The MIT Encyclopedia of Communication Disorders
The MIT Encyclopedia of Communication Disorders

Edited by Raymond D. Kent
Introduction ix
Acknowledgments xi

Part I: Voice 1
Acoustic Assessment of Voice 3
Aerodynamic Assessment of Vocal Function 7
Alaryngeal Voice and Speech Rehabilitation 10
Anatomy of the Human Larynx 13
Assessment of Functional Impact of Voice Disorders 20
Electroglostographic Assessment of Voice 23
Functional Voice Disorders 27
Hypokinetic Laryngeal Movement Disorders 30
Infectious Diseases and Inflammatory Conditions of the Larynx 32
Instrumental Assessment of Children’s Voice 35
Laryngeal Movement Disorders: Treatment with Botulinum Toxin 38
Laryngeal Reinnervation Procedures 41
Laryngeal Trauma and Peripheral Structural Ablations 45
Psychogenic Voice Disorders: Direct Therapy 49
The Singing Voice 51
Vocal Hygiene 54
Vocal Production System: Evolution 56
Vocalization, Neural Mechanisms of 59
Voice Acoustics 63
Voice Disorders in Children 67
Voice Disorders of Aging 72
Voice Production: Physics and Physiology 75
Voice Quality, Perceptual Evaluation of 78
Voice Rehabilitation After Conservation Laryngectomy 80
Voice Therapy: Breathing Exercises 82
Voice Therapy: Holistic Techniques 85
Voice Therapy for Adults 88
Voice Therapy for Neurological Aging-Related Voice Disorders 91
Voice Therapy for Professional Voice Users 95

Part II: Speech 99
Apraxia of Speech: Nature and Phenomenology 101
Apraxia of Speech: Treatment 104
Aprosodia 107
Augmentative and Alternative Communication Approaches in Adults 110
Augmentative and Alternative Communication Approaches in Children 112
Autism 115
Bilingualism, Speech Issues in 119
Developmental Apraxia of Speech 121
Dialect, Regional 124
Dysarthrias: Characteristics and Classification 126
Dysarthrias: Management 129
Dysphagia, Oral and Pharyngeal 132
Early Recurrent Otitis Media and Speech Development 135
Laryngectomy 137
Mental Retardation and Speech in Children 140
Motor Speech Involvement in Children 142
Mutism, Neurogenic 145
Orofacial Myofunctional Disorders in Children 147
Phonetic Transcription of Children’s Speech 150
Phonological Awareness Intervention for Children with Expressive Phonological Impairments 153
Phonological Errors, Residual 156
Phonology: Clinical Issues in Serving Speakers of African-American Vernacular English 158
Psychosocial Problems Associated with Communicative Disorders 161
Speech and Language Disorders in Children: Computer-Based Approaches 164
Speech and Language Issues in Children from Asian-Pacific Backgrounds 167
Speech Assessment, Instrumental 169
Speech Assessment in Children: Descriptive Linguistic Methods 174
Speech Development in Infants and Young Children with a Tracheostomy 176
Speech Disfluency and Stuttering in Children 180
Speech Disorders: Genetic Transmission 183
Speech Disorders in Adults, Psychogenic 186
Speech Disorders in Children: A Psycholinguistic Perspective 189
Speech Disorders in Children: Behavioral Approaches to Remediation 192
Speech Disorders in Children: Birth-Related Risk Factors 194
Speech Disorders in Children: Cross-Linguistic Data 196
Speech Disorders in Children: Descriptive Linguistic Approaches 198
Speech Disorders in Children: Motor Speech Disorders of Known Origin 200
Speech Disorders in Children: Speech-Language Approaches 204
Speech Disorders Secondary to Hearing Impairment Acquired in Adulthood 207
Speech Issues in Children from Latino Backgrounds 210
Speech Sampling, Articulation Tests, and Intelligibility in Children with Phonological Errors 213
Speech Sampling, Articulation Tests, and Intelligibility in Children with Residual Errors 215
Speech Sound Disorders in Children: Description and Classification 218
Stuttering 220
Transsexualism and Sex Reassignment: Speech Differences 223
Ventilator-Supported Speech Production 226

Part III: Language 229
Agrammatism 231
Agraphia 233
Alexia 236
Alzheimer’s Disease 240
Aphasia, Global 243
Aphasia, Primary Progressive 245
Aphasia: The Classical Syndromes 249
Aphasia, Wernicke’s 252
Aphasia Treatment: Computer-Aided Rehabilitation 254
Aphasia Treatment: Pharmacological Approaches 257
Aphasia Treatment: Psychosocial Issues 260
Aphasic Syndromes: Connectionist Models 262
Aphasiology, Comparative 265
Argument Structure: Representation and Processing 269
Attention and Language 272
Auditory-Motor Interaction in Speech and Language 275
Augmentative and Alternative Communication: General Issues 277
Bilingualism and Language Impairment 279
Communication Disorders in Adults: Functional Approaches to Aphasia 283
Communication Disorders in Infants and Toddlers 285
Communication Skills of People with Down Syndrome 288
Dementia 291
Dialect Speakers 294
Dialect Versus Disorder 297
Discourse 300
Discourse Impairments 302
Functional Brain Imaging 305
Inclusion Models for Children with Developmental Disabilities 307
Language Development in Children with Focal Lesions 311
Language Disorders in Adults: Subcortical Involvement 314
Language Disorders in African-American Children 318
Language Disorders in Latino Children 321
Language Disorders in School-Age Children: Aspects of Assessment 324
Language Disorders in School-Age Children: Overview 326
Language Impairment and Reading Disability 329
Language Impairment in Children: Cross-Linguistic Studies 331
Language in Children Who Stutter 333
Language of the Deaf: Acquisition of English 336
Language of the Deaf: Sign Language 339
Linguistic Aspects of Child Language Impairment—Prosody 344
Melodic Intonation Therapy 347
Memory and Processing Capacity 349
Mental Retardation 352
Morphosyntax and Syntax 354
Otis Media: Effects on Children’s Language 358
Perseveration 361
Phonological Analysis of Language Disorders in Aphasia 363
Phonology and Adult Aphasia 366
Poverty: Effects on Language 369
Pragmatics 372
Prelinguistic Communication Intervention for Children with Developmental Disabilities 375
Preschool Language Intervention 378
Prosodic Deficits 381
Reversibility/Mapping Disorders 383
Right Hemisphere Language and Communication Functions in Adults 386
Right Hemisphere Language Disorders 388
Segmentation of Spoken Language by Normal Adult Listeners 392
Semantics 395
Social Development and Language Impairment 398
Specific Language Impairment in Children 402
Syntactic Tree Pruning 405
Trace Deletion Hypothesis 407

Part IV: Hearing 411
Amplitude Compression in Hearing Aids 413
Assessment of and Intervention with Children Who Are Deaf or Hard of Hearing 421
Audition in Children, Development of 424
Auditory Brainstem Implant 427
Auditory Brainstem Response in Adults 429
Auditory Neuropathy in Children 433
Auditory Scene Analysis 437
Auditory Training 439
Classroom Acoustics 442
Cochlear Implants 447
Cochlear Implants in Adults: Candidacy 450
Cochlear Implants in Children 454
Dichotic Listening 458
Electrocochleography 461
Electronystagmography 467
Frequency Compression 471
Functional Hearing Loss in Children 475
Genetics and Craniofacial Anomalies 477
Hearing Aid Fitting: Evaluation of Outcomes 480
Hearing Aids: Prescriptive Fitting 482
Hearing Aids: Sound Quality 487
Hearing Loss and the Masking-Level Difference 489
Hearing Loss and Teratogenic Drugs or Chemicals 493
Hearing Loss Screening: The School-Age Child 495
Hearing Protection Devices 497
Masking 500
Middle Ear Assessment in the Child 504
Introduction

*The MIT Encyclopedia of Communication Disorders* (MITECD) is a comprehensive volume that presents essential information on communication sciences and disorders. The pertinent disorders are those that affect the production and comprehension of spoken language and include especially disorders of speech production and perception, language expression, language comprehension, voice, and hearing. Potential readers include clinical practitioners, students, and research specialists. Relatively few comprehensive books of similar design and purpose exist, so MITECD stands nearly alone as a resource for anyone interested in the broad field of communication disorders.

MITECD is organized into the four broad categories of Voice, Speech, Language, and Hearing. These categories represent the spectrum of topics that usually fall under the rubric of communication disorders (also known as speech-language pathology and audiology, among other names). For example, roughly these same categories were used by the National Institute on Deafness and Other Communication Disorders (NIDCD) in preparing its national strategic research plans over the past decade. The *Journal of Speech, Language, and Hearing Research*, one of the most comprehensive and influential periodicals in the field, uses the editorial categories of speech, language, and hearing. Although voice could be subsumed under speech, the two fields are large enough individually and sufficiently distinct that a separation is warranted. Voice is internationally recognized as a clinical and research specialty, and it is represented by journals dedicated to its domain (e.g., the *Journal of Voice*). The use of these four categories achieves a major categorization of knowledge but avoids a narrow fragmentation of the field at large. It is to be expected that the *Encyclopedia* would include cross-referencing within and across these four major categories. After all, they are integrated in the definitively human behavior of language, and disorders of communication frequently have wide-ranging effects on communication in its essential social, educational, and vocational roles.

In designing the content and structure of MITECD, it was decided that each of these major categories should be further subdivided into Basic Science, Disorders (nature and assessment), and Clinical Management (intervention issues). Although these categories are not always transparent in the entire collection of entries, they guided the delineation of chapters and the selection of contributors. These categories are defined as follows:

- **Basic Science** entries pertain to matters such as normal anatomy and physiology, physics, psychology and psychophysics, and linguistics. These topics are the foundation for clinical description and interpretation, covering basic principles and terminology pertaining to the communication sciences. Care was taken to avoid substantive overlap with previous MIT publications, especially the *MIT Encyclopedia of the Cognitive Sciences* (MITECS).
- **The Disorders** entries offer information on issues such as syndrome delineation, definition and characterization of specific disorders, and methods for the identification and assessment of disorders. As such, these chapters reflect contemporary nosology and nomenclature, as well as guidelines for clinical assessment and diagnosis.
- **The Clinical Management** entries discuss various interventions including behavioral, pharmacological, surgical, and prosthetic (mechanical and electronic). There is a general, but not necessarily one-to-one, correspondence between chapters in the Disorders and Clinical Management categories. For example, it is possible that several types of disorder are related to one general chapter on clinical management. It is certainly the case that different management strategies are preferred by different clinicians. The chapters avoid dogmatic statements regarding interventions of choice.

Because the approach to communicative disorders can be quite different for children and adults, a further cross-cutting division was made such that for many topics...
separate chapters for children and adults are included. Although some disorders that are first diagnosed in childhood may persist in some form throughout adulthood (e.g., stuttering, specific language impairment, and hearing loss may be lifelong conditions for some individuals), many disorders can have an onset either in childhood or in adulthood and the timing of onset can have implications for both assessment and intervention. For instance, when a child experiences a significant loss of hearing, the sensory deficit may greatly impair the learning of speech and language. But when a loss of the same degree has an onset in adulthood, the problem is not in acquiring speech and language, but rather in maintaining communication skills. Certainly, it is often true that an understanding of a given disorder has common features in both the developmental and acquired forms, but commonality cannot be assumed as a general condition.

Many decisions were made during the preparation of this volume. Some were easy, but others were not. In the main, entries are uniform in length and number of references. However, in a few instances, two or more entries were combined into a single longer entry. Perhaps inevitably in a project with so many contributors, a small number of entries were dropped because of personal issues, such as illness, that interfered with timely preparation of an entry. Happily, contributors showed great enthusiasm for this project, and their entries reflect an assembled expertise that is high tribute to the science and clinical practice in communication disorders.

Raymond D. Kent
MITECD began as a promising idea in a conversation with Amy Brand, a previous editor with MIT Press. The idea was further developed, refined, elaborated, and refined again in many ensuing e-mail communications, and I thank Amy for her constant support and assistance through the early phases of the project. When she left MIT Press, Tom Stone, Senior Editor of Cognitive Sciences, Linguistics, and Bradford Books, stepped in to provide timely advice and attention. I also thank Mary Avery, Acquisitions Assistant, for her help in keeping this project on track. I am indebted to all of them.

Speech, voice, language, and hearing are vast domains individually, and several associated editors helped to select topics for inclusion in MITECD and to identify contributors with the necessary expertise. The associate editors and their fields of responsibility are as follows:

Fred H. Bess, Ph.D., Hearing Disorders in Children
Joseph R. Duffy, Ph.D., Speech Disorders in Adults
Steven D. Gray, M.D. (deceased), Voice Disorders in Children
Robert E. Hillman, Ph.D., Voice Disorders in Adults
Sandra Gordon-Salant, Ph.D., Hearing Disorders in Adults
Mabel L. Rice, Ph.D., Language Disorders in Children
Lawrence D. Shriberg, Ph.D., Speech Disorders in Children
David A. Swinney, Ph.D., and Lewis P. Shapiro, Ph.D., Language Disorders in Adults

The advice and cooperation of these individuals is gratefully acknowledged. Sadly, Dr. Steven D. Gray died within the past year. He was an extraordinary man, and although I knew him only briefly, I was deeply impressed by his passion for knowledge and life. He will be remembered as an excellent physician, creative scientist, and valued friend and colleague to many.

Dr. Houri Vorperian greatly facilitated this project through her inspired planning of a computer-based system for contributor communications and record management. Sara Stuntebeck and Sara Brost worked skillfully and accurately on a variety of tasks that went into different phases of MITECD. They offered vital help with communications, file management, proofreading, and the various and sundry tasks that stood between the initial conception of MITECD and the submission of a full manuscript.

P. M. Gordon and Associates took on the formidable task of assembling 200 entries into a volume that looks and reads like an encyclopedia. I thank Denise Bracken for exacting attention to the editing craft, creative solutions to unexpected problems, and forbearance through it all.

MITECD came to reality through the efforts of a large number of contributors—too many for me to acknowledge personally here. However, I draw the reader’s attention to the list of contributors included in this volume. I feel a sense of community with all of them, because they believed in the project and worked toward its completion by preparing entries of high quality. I salute them not only for their contributions to MITECD but also for their many career contributions that define them as experts in the field. I am honored by their participation and their patient cooperation with the editorial process.

Raymond D. Kent
Part I: Voice
Acoustic Assessment of Voice

Acoustic assessment of voice in clinical applications is dominated by measures of fundamental frequency ($f_0$), cycle-to-cycle perturbations of period (jitter) and intensity (shimmer), and other measures of irregularity, such as noise-to-harmonics ratio (NHR). These measures are widely used, in part because of the availability of electronic and microcomputer-based instruments (e.g., Kay Elemetrics Computerized Speech Laboratory [CSL] or Multispeech, Real-Time Pitch, Multi-Dimensional Voice Program [MDVP], and other software/hardware systems), and in part because of long-term precedent for perturbation (Lieberman, 1961) and spectral noise measurements (Yanagihara, 1967). Absolute measures of vocal intensity are equally basic but require calibrations and associated instrumentation (Winholz and Titze, 1997).

Independently, these basic acoustic descriptors—$f_0$, intensity, jitter, shimmer, and NHR—can provide some very basic characterizations of vocal health. The first two, $f_0$ and intensity, have very clear perceptual correlates—pitch and loudness, respectively—and should be assessed for both stability and variability and compared to age and sex norms (Kent, 1994; Baken and Orlikoff, 2000). Ideally, these tasks are recorded over headset microphones with direct digital acquisition at very high sampling rates (at least 48 kHz). The materials to be assessed should be obtained following standardized elicitation protocols that include sustained vowel phonations at habitual levels, levels spanning a client’s vocal range in both $f_0$ and intensity, running speech, and speech tasks designed to elicit variation (Titze, 1995; Awan, 2001). Note, however, that not all measures will be appropriate for all tasks; perturbation statistics, for example, are usually valid only when extracted from sustained vowel phonations.

These basic descriptors are not in any way comprehensive of the range of available measures or the available signal properties and dimensions. Table 1 categorizes measures (Buder, 2000) based on primary basic signal representations from which measures are derived. Although these categories are intended to be exhaustive and mutually exclusive, some more modern algorithms process components through several types. (For more detail on the measurement types, see Buder, 2000, and Baken and Orlikoff, 2000.) Modern algorithmic approaches should be selected for (1) interpretability with respect to aerodynamic and physiological models of phonation and (2) the incorporation of multivariate measures to characterize vocal function.

Interdependence of Basic Measures. The interdependence between $f_0$ and intensity is mapped in a voice range profile, or phonetogram, which is an especially valuable assessment for the professional voice user (Coleman, 1993). Furthermore, the dependence of perturbations and signal-to-noise ratios on both $f_0$ and intensity is well known (Klingholz, 1990; Pabon, 1991). This dependence is not often assessed rigorously, perhaps because of the time-consuming and strenuous nature of a full voice profile. However, an abbreviated or focused profiling in which samples related to habitual $f_0$ by a set number of semitones, or related to habitual intensity by a set number of decibels, could be standardized to control for this dependence efficiently. Finally, it should be understood that perturbations and NHR-type measures will usually covary for many reasons, the simplest ones being methodological (Hillenbrand, 1987): an increase in any one of the underlying phenomena detected by a single measure will also affect the other measures.

**Table 1. Outline of Traditional Acoustic Algorithm Types**

<table>
<thead>
<tr>
<th>f&lt;sub&gt;0&lt;/sub&gt; statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term perturbations</td>
</tr>
<tr>
<td>Long-term perturbations</td>
</tr>
<tr>
<td>Amplitude statistics</td>
</tr>
<tr>
<td>Short-term perturbations</td>
</tr>
<tr>
<td>Long-term perturbations</td>
</tr>
<tr>
<td>f&lt;sub&gt;0&lt;/sub&gt;/amplitude covariations</td>
</tr>
<tr>
<td>Waveform perturbations</td>
</tr>
<tr>
<td>Spectral measures</td>
</tr>
<tr>
<td>Spectrogram measures</td>
</tr>
<tr>
<td>Fourier and LPC spectra</td>
</tr>
<tr>
<td>Long-term average spectra</td>
</tr>
<tr>
<td>Cepstra</td>
</tr>
<tr>
<td>Inverse filter measures</td>
</tr>
<tr>
<td>Radiated signal</td>
</tr>
<tr>
<td>Flow-mask signals</td>
</tr>
<tr>
<td>Dynamic measures</td>
</tr>
</tbody>
</table>

Periodicity as a Reference. The chief problem with nearly all acoustic assessments of voice is the determination of $f_0$. Most voice quality algorithms are based on the prior identification of the periodic component in the signal (based on glottal pulses in the time domain or harmonic structure in the frequency domain). Because phonation is ideally a nearly periodic process, it is logical to conceive of voice measures in terms of the degree to which a given sample deviates from pure periodicity. There are many conceptual problems with this simplification, however. At the physiological level, glottal morphology is multidimensional—superior-inferior asymmetry is a basic feature of the two-mass model (Ishizaka and Flanagan, 1972), and some anterior-posterior asymmetry is also inevitable—rendering it unlikely that a glottal pulse will be marked by a discrete or even a single instant of glottal closure. At the level of the signal, the deviations from periodicity may be either random or correlated, and in many cases they are so extreme as to preclude identification of a regular period. Finally, at the perceptual level, many factors related to deviations from a pure $f_0$ can contribute to pitch perception (Zwicker and Fastl, 1990).

At any or all of these levels, it becomes questionable to characterize deviations with pure periodicity as a reference. In acoustic assessment, the primary level of concern is the signal. The National Center for Voice and
Speech issued a summary statement (Titze, 1995) recommending a typology for categorizing deviations from periodicity in voices (see also Baken and Orlikoff, 2000, for further subtypes). This typology capitalizes on the categorical nature of dynamic states in nonlinear systems; all the major categories, including stable points, limit cycles, period-doubling/tripling/. . . , and chaos can be observed in voice signals (Herzel et al., 1994; Sataloff and Hawkshaw, 2001). As in most highly nonlinear dynamic systems, deviations from periodicity can be categorized on the basis of bifurcations, or sudden qualitative changes in vibratory pattern from one of these states to another.

Figure 1 displays a common form for one such bifurcation and illustrates the importance of accounting for its presence in the application of perturbation measures. In this sustained vowel phonation by a middle-aged woman with spasmodic dysphonia, a transition to subharmonics is clearly visible in segment b (similar patterns occur in individuals without dysphonias). Two $f_0$ extractions are presented for this segment, one at the targeted level of approximately 250 Hz and another which the tracker finds one octave below this; inspection of the waveform and a perceived biphonia both justify this 125-Hz analysis as a new fundamental frequency, although it can also be understood in this context as a subharmonic to the original fundamental. There is therefore some ambiguity as to which fundamental is valid during this episode, and an automatic analysis could plausibly identify either frequency. (Here the waveform-matching algorithm implemented in CSpeechSP [Milenkovic, 1997] does identify either frequency, depending on where in the waveform the algorithm is applied; initiating the algorithm within the subharmonic segment predisposes it to identify the lower fundamental.)

The acoustic measures of the segments displayed in Figure 1 reveal the nontrivial differences that result, depending on the basic glottal pulse form under consideration. When the pulses of segment a are considered,
the perturbations around the base period associated with the high $f_0$ are low and normative; in segment $b$, perturbations around the longer periods of the lower $f_0$ are still low (jitter is improved, while shimmer and the signal-to-noise ratio show some degradation). However, when all segments are considered together to include the perturbations around the high $f_0$ tracked through segment $b$ and into $c$, the perturbation statistics are all increased by an order of magnitude. Many important methodological and theoretical questions should be raised by such common scenarios in which we must consider not just voice typing, but the segment-by-segment validity of applying perturbation measures with a particular $f_0$ as reference. If, as is often assumed, jitter and shimmer are ascribed to “random” variations, then the correlated modulations of a strong subharmonic episode should be excluded. Alternatively, the perturbations might be analyzed with respect to the subharmonic $f_0$. In any case, assessment by means of perturbation statistics with no consideration of their underlying sources is unwise.

**Perceptual, Aerodynamic, and Physiological Correlates of Acoustic Measures.** Regarding perceptual voice ratings, Gerratt and Kreiman (2000) have critiqued traditional assessments on several important methodological and theoretical points. However, these points may not apply to acoustic analysis if (1) acoustic analysis is validated on its own success and not exclusively in relation to the problematic perceptual classifications, and (2) acoustic analysis is thoroughly grounded for interpretation in some clear aerodynamic or physiological model of phonation. Gerratt and Kreiman also argue that clinical classification may not be derived along a continuum that is defined with reference to normal qualities, but again, this argument may need to be reversed for the acoustic domain. It is only by reference to a specific model that any assessment on acoustic grounds can be interpreted (though this does not preclude development of an independent model for a pathological phonatory mechanism). In clinical settings, acoustic voice assessment often serves to corroborate perceptual assessment. However, as guided by auditory experience and in conjunction with the ear and other instrumental assessments, careful acoustic analysis can be oriented to the identification of physiological status.

In attempting to draw safe and reasonably direct inferences from acoustic signal, aerodynamic models correlated with high-frequency harmonic energy), and (e) pulse skewing (which is negatively correlated with low-frequency harmonic energy; this low-frequency region is also positively correlated with open quotient and peak volume velocity measures of the glottal flow waveform). The effect of turbulence due to high airflow through the glottis is schematized by (d), indicating the associated appearance of high-frequency aperiodic energy in the spectrum. See voice acoustics for other graphical and quantitative associations between glottal status and spectral characteristics.

![Figure 2. Spectral features associated with models of phonation, including the Liljencrants-Fant (LF) model of glottal flow and aperiodicity source models developed by Stevens. The LF model of glottal flow is shown at top left. At bottom left is the LF model of glottal flow derivative, showing the rate of change in flow. At right is a spectrum schematic showing four effects. These effects include three derived parameters of the LF model: (a) excitation strength (the maximum negative amplitude of the flow derivative, which is positively correlated with overall harmonic energy), (b) dynamic leakage or non-zero return phase following the point of maximum excitation (which is negatively correlated with high-frequency harmonic energy), and (c) pulse skewing (which is negatively correlated with low-frequency harmonic energy; this low-frequency region is also positively correlated with open quotient and peak volume velocity measures of the glottal flow waveform). The effect of turbulence due to high airflow through the glottis is schematized by (d), indicating the associated appearance of high-frequency aperiodic energy in the spectrum. See voice acoustics for other graphical and quantitative associations between glottal status and spectral characteristics.](image-url)
of glottal behavior present important links to the physiological domain. Attempts to recover the glottal flow waveform, either from a face mask-transduced flow recording (Rothenberg, 1973) or a microphone-transduced acoustic recording (Davis, 1975), have proved to be labor-intensive and prone to error (Ni Chasaide and Gobl, 1997). Rather than attempting to eliminate the effects of the vocal tract, it may be more fruitful to understand its in situ relationship with phonation, and infer, via the types of features displayed in Figure 2, the status of the glottis as a sound source. Interpretation of spectral features, such as the amplitudes of the first harmonics and at the formant frequencies, may be an effective alternative when guided by knowledge of glottal aerodynamics and acoustics (Hanson, 1997; Ni Chasaide and Gobl, 1997; Hanson and Chuang, 1999). Deep familiarity with acoustic mechanisms is essential for such interpretations (Titze, 1994; Stevens, 1998), as is a model with clear and meaningful parameters, such as the Liljencrants-Fant (LF) model (Fant, Liljencrants, and Lin, 1985). These parameters of the LF model have proved to be meaningful in acoustic studies (Gauflin & Sundberg, 1989) and useful in refined efforts at inverse filtering (Fröhlich, Michaelis, and Strube, 2001). Figure 2 summarizes selected parameters of the LF source model following Ni Chasaide and Gobl (1997) and the glottal turbulence source following Stevens (1998); see also VOICE ACOUSTICS for other approaches relating glottal status to spectral measures.

Other spectral-based measures implement similar model-based strategies by selecting spectral component ratios (e.g., the VTI and SPI parameters of MDVP). Sophisticated spectral noise characterizations control for perturbations and modulations (Murphy, 1999; Qi, Hillman, and Milstein, 1999), or employ curve-fitting and statistical models to produce more robust measures (Alku, Strik, and Vilkmn, 1997; Michaelis, Fröhlich, and Strube, 1998; Schoentgen, Bensaid, and Bucella, 2000). A particularly valuable modern technique for detecting turbulence at the glottis, the glottal-to-noise-excitation ratio (Michaelis, Gramss, and Strube, 1997), has been especially successful in combination with other measures (Fröhlich et al., 2000). The use of acoustic techniques for voice will only improve with the inclusion of more knowledge-based measures in multivariate representations (Wolfe, Cornell, and Palmer, 1991; Callen et al., 2000; Wuyts et al., 2000).

—Eugene H. Buder

References


Aerodynamic Assessment of Vocal Function

A number of methods have been used to quantitatively assess the air volumes, airflows, and air pressures involved in voice production. The methods have been mostly used in research to investigate mechanisms that underlie normal and disordered voice and speech production. The clinical use of aerodynamic measures to assess patients with voice disorders has been increasing (Colton and Casper, 1996; Hillman, Montgomery, and Zeitels, 1997; Hillman and Kobler, 2000).

Measurement of Air Volumes. Respiratory research in human communication has focused primarily on the measurement of the air volumes that are typically expended during selected speech and singing tasks, and on specifying the ranges of lung inflation levels across which such tasks are normally performed (cf. Hixon, Goldman, and Mead, 1973; Watson and Hixon, 1985; Hoit and Hixon, 1987; Hoit et al., 1990). Air volumes are measured in standard metric units (liters, cubic centimeters, milliliters) and lung inflation levels are usually specified in terms of a percentage of the vital capacity or total lung volume.

Both direct and indirect methods have been used to measure air volumes expended during phonation. Direct measurement of orally displaced air volumes during phonatory tasks can be accomplished, to a limited extent, by means of a mouthpiece or face mask connected to a measurement device such as a spirometer (Beckett, 1971) or pneumotachograph (Isshiki, 1964). The use of a mouthpiece essentially limits speech production to sustained vowels, which are sufficient for assessing selected volumetric-based phonatory parameters. There are also concerns that face masks interfere with normal jaw movements and that the oral acoustic signal is degraded, so that auditory feedback is reduced or distorted and simultaneous acoustic analysis is limited. These limitations, which are inherent to the use of devices placed in or around the mouth to directly collect oral airflow, plus additional measurement-related restrictions (Hillman and Kobler, 2000) have helped motivate the development and application of indirect measurement approaches.

Most speech breathing research has been carried out using indirect approaches for estimating lung volumes by means of monitoring changes in body dimensions. The basic assumption underlying the indirect approaches is that changes in lung volume are reflected in proportional changes in body torso size. One relatively cumbersome but time-honored approach has been to place subjects in a sealed chamber called a body plethysmograph to allow estimation of the air volume displaced by the body during respiration (Draper, Ladefoged, and Whitteridge, 1959). More often used for speech breathing research are transducers (magnetometers: Hixon, Goldman, and Mead, 1973; inductance plethysmographs: Sperry, Hillman, and Perkell, 1994) that unobtrusively monitor changes in the dimensions of the rib cage and abdomen (referred to collectively as the chest wall) that account for the majority of respiratory-related changes in torso dimension (Mead et al., 1967). These approaches have been primarily employed to study respiratory function during continuous speech and singing tasks that include both voiced and voiceless sound production, as opposed to assessing air volume usage during phonatory tasks that involve only laryngeal production of voice (e.g., sustained vowels). There are also ongoing efforts to develop more accurate methods for non-invasively monitoring chest wall activity to capture finer details of how the three-dimensional geometry of the body is altered during respiration (see Cala et al., 1996).

Measurement of Airflow. Airflow associated with phonation is usually specified in terms of volume velocity (i.e., volume of air displaced per unit of time). Volume velocity airflow rates for voice production are typically reported in metric units of volume displaced (liters or cubic centimeters) per second.
Estimates of average airflow rates can be obtained by simply dividing air volume estimates by the duration of the phonatory task. Average glottal airflow rates have usually been estimated during vowel phonation by using a mouthpiece or face mask to channel the oral air stream through a pneumotachograph (Isshiki, 1964). There has also been somewhat limited use of hot wire anemometer devices (mounted in a mouthpiece) to estimate average glottal airflow during sustained vowel phonation (Woo, Colton, and Shangold, 1987). Estimates of average glottal airflow rates can be obtained from the oral airflow during vowel production because the vocal tract is relatively nonconstricted, with no major sources of turbulent airflow between the glottis and the lips.

There have also been efforts to obtain estimates of the actual airflow waveform that is generated as the glottis rapidly opens and closes during flow-induced vibration of the vocal folds (the glottal volume velocity waveform). The glottal volume velocity waveform cannot be directly observed by measuring the oral airflow signal because the waveform is highly convoluted by the resonance activity (formants) of the vocal tract. Thus, recovery of the glottal volume velocity waveform requires methods that eliminate or correct for the influences of the vocal tract. This has typically been accomplished aerodynamically by processing the output of a fast-responding pneumotachograph (high-frequency response) using a technique called inverse filtering, in which the major resonances of the vocal tract are estimated and the oral airflow signal is processed (inverse filtered) to eliminate them (Rothenberg, 1977; Holmberg, Hillman, and Perkell, 1988).

![Figure 1. Instrumentation and resulting signals for simultaneous collection of oral airflow, intraoral air pressure, the acoustic signal, and chest wall (rib cage and abdomen) dimensions during production of the syllable string /pi-pi-pi/. Signals shown in the bottom panel are processed and measured to provide estimates of average glottal airflow rate, average subglottal air pressure, lung volume, and glottal waveform parameters.](image-url)
Measurement of Air Pressure. Measurements of air pressures below (subglottal) and above (supraglottal) the vocal folds are of primary interest for characterizing the pressure differential that must be achieved to initiate and maintain vocal fold vibration during normal exhalatory phonation. In practice, air pressure measurements related specifically to voice production are typically acquired during vowel phonation when there are no vocal tract constrictions of sufficient magnitude to build up positive supraglottal pressures. Under these conditions, it is usually assumed that supraglottal pressure is essentially equal to atmospheric pressure and only subglottal pressure measurements are obtained. Air pressures associated with voice and speech production are usually specified in centimeters of water (cm H₂O).

Both direct and indirect methods have been used to measure subglottal air pressures during phonation. Direct measures of subglottal air pressure can be obtained by inserting a hypodermic needle into the subglottal airway through a puncture in the anterior neck at the cricothyroid space (Isshiki, 1964). The needle is connected to a pressure transducer by tubing. This method is very accurate but also very invasive. It is also possible to insert a very thin catheter through the posterior cartilaginous glottis (between the arytenoids) to sense subglottal air pressure during phonation, or to use an array of miniature transducers positioned directly above and below the glottis (Cranen and Boves, 1985). These methods cannot be tolerated by all subjects, and the heavy topical anesthetization of the larynx that is required can affect normal function.

Indirect estimates of tracheal (subglottal) air pressure can be obtained via the placement of an elongated balloon-like device into the esophagus (Liberman, 1968). The deflated esophageal balloon is attached to a catheter that is typically inserted transnasally and then swallowed into the esophagus to be positioned at the midthoracic level. The catheter is connected to a pressure transducer and the balloon is slightly inflated. Accurate use of this invasive method also requires simultaneous monitoring of lung volume.

Noninvasive, indirect estimates of subglottal air pressure can be obtained by measuring intraoral air pressure during specially constrained utterances (Smitheran and Hixon, 1981). This is usually done by sensing air pressure just behind the lips with a transtabulary placed catheter connected to a pressure transducer. These intraoral pressure measures are obtained as subjects produce strings of bilabial /p/ + vowel syllables (e.g., /pi-pi-pi-pi-pi/) at constant pitch and loudness. This method works because the vocal folds are abducted during /p/ production, thus allowing pressure to equilibrate throughout the airway, making intraoral pressure equal to subglottal pressure (Fig. 1).

Additional Derived Measures. There have been numerous attempts to extend the utility of aerodynamic measures by using them in the derivation of additional parameters aimed at better elucidating underlying mechanisms of vocal function. Such derived measures usually take the form of ratios that relate aerodynamic parameters to each other, or that relate aerodynamic parameters to simultaneously obtained acoustic measures. Common examples include (1) airflow (glottal) resistance (see Smitheran and Hixon, 1981), (2) vocal efficiency (Schutte, 1980; Holmberg, Hillman, and Perkell, 1988), and (3) measures that interrelate glottal volume velocity waveform parameters (Holmberg, Hillman, and Perkell, 1988).

Normative Data. As is the case for most measures of vocal function, there is not currently a set of normative data for aerodynamic measures that is universally accepted and applied in research and clinical work. Methods for collecting such data have not been standardized, and study samples have generally not been of sufficient size or appropriately stratified in terms of age and sex to ensure unbiased estimates of underlying aerodynamic phonatory parameters in the normal population. However, there are several sources in the literature that provide estimates of normative values for selected aerodynamic measures (Kent, 1994; Baken, 1996; Colton and Casper, 1996).

See also VOICE PRODUCTION: PHYSICS AND PHYSIOLOGY.

—Robert E. Hillman

References


Alaryngeal Voice and Speech Rehabilitation

Loss of the larynx due to disease or injury will result in numerous and significant changes that cross anatomical, physiological, psychological, social, psychosocial, and communication domains. Surgical removal of the larynx, or total laryngectomy, involves resectioning the entire framework of the larynx. Although total laryngectomy may occur in some instances due to traumatic injury, the majority of cases worldwide are the result of cancer. Approximately 75% of all laryngeal tumors arise from squamous epithelial tissue of the true vocal fold (Bailey, 1985). In some instances, and because of the location of many of these lesions, less aggressive approaches to medical intervention may be pursued. This may include radiation therapy or partial surgical resection, which seeks to conserve portions of the larynx, or the use of combined chemoradiation protocols (Hillman et al., 1998; Orlikoff et al., 1999). However, when malignant lesions are sufficiently large or when the location of the tumor threatens the lymphatic compartment of the larynx, total laryngectomy is often indicated for reasons of oncological safety (Doyle, 1994).

Effects of Total Laryngectomy

The two most prominent effects of total laryngectomy as a surgical procedure are change of the normal airway and loss of the normal voicing mechanism for verbal communication. Once the larynx is surgically removed from the top of the trachea, the trachea is brought forward to the anterior midline neck and sutured into place near the sternal notch. Thus, total laryngectomy necessitates that the airway be permanently separated from the upper aerodynamic (oral and pharyngeal) pathway. When the laryngectomy is completed, the tracheal airway will remain separate from the oral cavity, pharynx, and esophagus. Under these circumstances, not only is the primary structure for voice generation lost, but the intimate relationship between the pulmonary system and that of the structures of the upper airway, and consequently the vocal tract, is disrupted. Therefore, if verbal communication is to be acquired and used post-laryngectomy, an alternative method of creating an alaryngeal voice source must be achieved.

Methods of Postlaryngectomy Communication

Following laryngectomy, the most significant communicative component to be addressed via voice and speech rehabilitation is the lost voice source. Once the larynx is removed, some alternative method of providing a new, “alaryngeal” sound source is required. There are two general categories in which an alternative, alaryngeal voice source may be achieved. These categories are best described as intrinsic and extrinsic methods. The distinction between these two methods is contingent on the manner in which the alaryngeal voice source is achieved. Intrinsic alaryngeal methods imply that the alaryngeal voice source is found within the system; that is, alternative physical-anatomical structures are used to generate sound. In contrast, extrinsic methods of alaryngeal speech rely on the use of an external sound source, typically an electronic source, or what is termed the artificial larynx, or the electrolarynx. The fundamental differences between intrinsic and extrinsic methods of alaryngeal speech are discussed below.

Intrinsic Methods of Alaryngeal Speech

The two most prominent methods of intrinsic alaryngeal speech are esophageal speech (Diedrich, 1966; Doyle, 1994) and tracheoesophageal (TE) speech (Singer and Blom, 1980). While these two intrinsic methods of alaryngeal speech are dissimilar in some respects, both rely on generation of an alaryngeal voice source by creating oscillation of tissues in the area of the lower pharynx and upper esophagus. This vibratory structure is somewhat variable in regard to width, height, and location (Diedrich and Youngstrom, 1966; Damste, 1986); hence, the preferred term for this alaryngeal voicing source is the pharyngoesophageal (PE) segment. One
muscle that comprises the PE segment is the cricopharyngeal muscle. Beyond the commonality in the use of the PE segment as a vicarious voicing source for both esophageal and TE methods of alaryngeal speech, the manner in which these methods are achieved does differ.

**Esophageal Speech.** For esophageal speech, the speaker must move air from the oral cavity across the tonically closed PE segment in order to insufflate the esophageal reservoir (located inferior to the PE segment). Two methods of insufflation may be utilized. These methods might be best described as being either direct or indirect approaches to insufflation. Direct methods require the individual speaker to actively manipulate air in the oral cavity to effect a change in pressure. When pressure build-up is achieved in the oral cavity via compression maneuvers, and when the pressure becomes of sufficient magnitude to overcome the muscular resistance of the PE segment, air will move across the segment (inferiorly) into the esophagus. This may be accomplished with nonspeech tasks (tongue maneuvers) or as a result of producing specific sounds (e.g., stop consonants).

In contrast, for the indirect (inhalation) method of air insufflation, the speaker indirectly creates a negative pressure in the esophageal reservoir via rapid inhalation through the tracheostoma. This results in a negative pressure in the esophagus relative to the normal atmospheric pressure within the oral cavity/vocal tract (Diedrich and Youngstrom, 1966; Diedrich, 1968; Doyle, 1994). Air then moves passively across the PE segment in order to equalize pressures between the pharynx and esophagus. Once insufflation occurs, this air can be used to generate PE segment vibration in the same manner following other methods of air insufflation. While a distinction between direct and indirect methods permits increased understanding of the physical requirements for esophageal voice production, many esophageal speakers who exhibit high levels of proficiency will often utilize both methods for insufflation. Regardless of which method of air insufflation is used, this air can then be forced back up across the PE segment, and as a result, the tissue of this sphincter will oscillate. This esophageal sound source can then be manipulated in the upper regions of the vocal tract into the sounds of speech.

The acquisition of esophageal speech is a complex process of skill building that must be achieved under the direction of an experienced instructor. Clinical emphasis typically involves tasks that address four skills believed to be fundamental to functional esophageal speech (Berlin, 1963): (1) the ability to phonate reliably on demand, (2) the ability to maintain a short latency between air insufflation and esophageal phonation, (3) the ability to maintain adequate duration of voicing, and (4) the ability to sustain voicing while articulating. These foundational skills have been shown to reflect those progressive abilities that have historically defined speech skills of “superior” esophageal speakers (Wepman et al., 1953; Snidecor, 1968). However, the successful acquisition of esophageal speech may be limited, for many reasons.

Regardless of which method of insufflation is used, esophageal speakers will exhibit limitations in the physical dimensions of speech. Specifically, fundamental frequency is reduced by about one octave (Curry and Snidecor, 1961), intensity is reduced by about 10 dB SPL from that of the normal speaker (Weinberg, Hori, and Smith, 1980), and the durational characteristics of speech are also reduced. Speech intelligibility is also decreased due to limits in the aerodynamic and voicing characteristics of esophageal speech. As it is not an abductory-adductory system, voiced-for-voiceless perceptual errors (e.g., perceptual identification of b for p) are common. This is a direct consequence of the esophageal speaker’s inability to insufflate large or continuous volumes of air into the reservoir. Esophageal speakers must frequently reinsufflate the esophageal reservoir to maintain voicing. Because of this, it is not uncommon to see esophageal speakers exhibit pauses at unusual points in an utterance, which ultimately alters the normal rhythm of speech. Similarly, the prosodic contour of esophageal speech and associated features is often perceived to be abnormal. In contrast to esophageal speech, the TE method capitalizes on the individual’s access to pulmonary air for esophageal insufflation, which offers several distinct advantages relative to esophageal speech.

**Tracheoesophageal Speech.** TE speech uses the same voicing source as traditional esophageal speech, the PE segment. However, in TE speech the speaker is able to access and use pulmonary air as a driving source. This is achieved by the surgical creation of a controlled midline puncture in the trachea, followed by insertion of a one-way TE puncture voice prosthesis (Singer and Blom, 1980), either at the time of laryngectomy or as a second procedure at some point following laryngectomy. Thus, TE speech is best described as a surgical-prosthetic method of voice restoration. Though widely used, TE voice restoration is not problem-free. Limitations in application must be considered, and complications may occur.

The design of the TE puncture voice prosthesis is such that when the tracheostoma is occluded, either by hand or via use of a complementary tracheostoma breathing valve, air is directed from the trachea through the prosthesis and into the esophageal reservoir. This access permits a variety of frequency, intensity, and durational variables to be altered in a fashion different from that of the traditional esophageal speaker (Robbins et al., 1984; Pauloski, 1998). Because the TE speaker has direct access to a pulmonary air source, his or her ability to modify the physical (frequency, intensity, and durational) characteristics of the signal in response to changes in the aerodynamic driving source, along with associated changes in prosodic elements of the speech signal (i.e., stress, intonation, juncture), is enhanced considerably. Such changes have a positive impact on auditory-perceptual judgments of this method of alaryngeal speech.

While the frequency of TE speech is still reduced from that of normal speech, the intensity is greater, and the
durational capabilities meet or exceed those of normal speakers (Robbins et al., 1984). Finally, research into the influence of increased aerodynamic support in TE speakers relative to traditional esophageal speech on speech intelligibility has suggested that positive effects may be observed (Doyle, Danhauer, and Reed, 1988) despite continued voiced-for-voiceless perceptual errors. Clearly, the rapidity of speech reacquisition in addition to the relative increases in speech intelligibility and the changes in the overall physical character of TE speech offers considerable advantages from the perspective of communication rehabilitation.

Artificial Laryngeal Speech. Extrinsic methods of alaryngeal voice production are common. Although some pneumatic devices have been introduced, they are not widely used today. The most frequently used extrinsic method of producing alaryngeal speech uses an electronic artificial larynx, or electrolarynx. These devices provide an external energy (voice) source that is introduced either directly into the oral cavity (intraoral) or by placing a device directly on the tissues of the neck (transcervical). Whether the electrolaryngeal tone is introduced into the oral cavity directly or through transmission via tissues of the neck, the speaker is able to modulate the electrolaryngeal source into speech.

The electrolarynx is generally easy to use. Speech can be acquired relatively quickly, and the device offers a reasonable method of functional communication to those who have undergone total laryngectomy (Doyle, 1994). Its major limitations have traditionally related to negative judgments of electrolaryngeal speech relative to the mechanical nature of many devices. Current research is seeking to modify the nature of the electronic sound source produced. The intelligibility of electrolaryngeal speech is relatively good, given the external nature of the alaryngeal voice source and the electronic character of sound production. A reduction in speech intelligibility is primarily observed for voiceless consonants (i.e., voiceless errors) due to the fact that the electrolarynx is a continuous sound source (Weiss and Basili, 1985).

Rehabilitative Considerations

All methods of alaryngeal speech, whether esophageal, TE, or electrolaryngeal, have distinct advantages and disadvantages. Advantages for esophageal speech include a nonmechanical and hands-free method of communication. For TE speech, pitch is near normal, loudness exceeds normal, and speech rate and prosody is near normal; for artificial larynx speech, it may be acquired quickly by most people and may be used in conditions of background noise. In contrast, disadvantages for esophageal speech include lowered pitch, loudness, and speech rate. For TE speech, it involves use and maintenance of a prosthetic device with associated costs; for artificial larynx speech, a mechanical quality is common and it requires the use of one hand. While “normal” speech cannot be restored with these methods, no matter how proficient the speaker’s skills, all methods are viable postlaryngectomy communication options, and at least one method can be used with a functional communicative outcome in most instances. Professionals who work with individuals who have undergone total laryngectomy must focus on identifying a method that meets each speaker’s particular needs. Although clinical intervention must focus on making any given alaryngeal method as proficient as possible, the individual speaker’s needs, as well as the relative strengths and weaknesses of each method, must be considered. In this way, use of a given method may be enhanced so that the individual may achieve the best level of social reentry following laryngectomy. Further, nothing prevents an individual from using multiple methods of alaryngeal speech, although one or another may be preferred in a given communication context or environment. But an important caveat is necessary: Just because a method of alaryngeal speech has been acquired and it has been deemed “proficient” at the clinical level (e.g., results in good speech intelligibility) and is “functional” for basic communication purposes, this does not imply that “rehabilitation” has been successfully achieved.

The reacquisition of verbal communication is without question a critical component of recovery and rehabilitation postlaryngectomy; however, it is only one dimension of the complex picture of a successful return to a normal life as possible. All individuals who have undergone a laryngectomy will confront myriad restrictions in multiple domains, including anatomical, physiological, psychological, communicative, and social. As a result, postlaryngectomy rehabilitation efforts that address these areas may increase the likelihood of a successful postlaryngectomy outcome.

See also LARYNGECTOMY.

—Philip C. Doyle and Tanya L. Eadie

References


Further Readings


**Anatomy of the Human Larynx**

The larynx is an organ that sits in the hypopharynx, at the crossroads of the upper respiratory and upper digestive tracts. The larynx is intimately involved in respiration, deglutition, and phonation. Although it is the primary sound generator of the peripheral speech mechanism, it must be viewed primarily as a respiratory organ. In this capacity it controls the flow of air into and out of the lower respiratory tract, prevents food from becoming lodged in the trachea or bronchi (which would threaten life and interfere with breathing), and, through
the cough reflex, assists in dislodging material from the lower airway. The larynx also plays a central role in the development of the intrathoracic and intra-abdominal pressures needed for lifting, elimination of bodily wastes, and sound production.

Throughout life, the larynx undergoes maturational and involutional (aging) changes (Kahane, 1996), which influence its capacity as a sound source. Despite these naturally and slowly occurring structural changes, the larynx continues to function relatively flawlessly. This is a tribute to the elegance of its structure.

**Regional Anatomical Relationships.** The larynx is located in the midline of the neck. It lies in front of the vertebral column and between the hyoid bone above and the trachea below. In adults, it lies between the third and sixth cervical vertebrae. The root, or pharyngeal portion, of the tongue is interconnected with the epiglottis of the larynx by three fibroelastic bands, the glossoepiglottic folds. The lowermost portion of the pharynx, the hypopharynx, surrounds the posterior aspect of the larynx. Muscle fibers of the inferior pharyngeal constrictor attach to the posterolateral aspect of the thyroid and coid cartilages. The esophagus lies inferior and posterior to the larynx. It is a muscular tube that interconnects the pharynx and the stomach. Muscle fibers originating from the coid cartilage form part of the muscular valve, which opens to allow food to pass from the pharynx into the esophagus.

**Cartilaginous Skeleton.** The larynx is composed of five major cartilages: thyroid, coid, one pair of arytenoids, and the epiglottis (Fig. 1). The hyoid bone, though intimately associated with the larynx, is not part of it. The cartilaginous components of the larynx are joined by ligaments and membranes. The thyroid and coid cartilages are composed of hyaline cartilage, which provides them with form and rigidity. They are interconnected by the cricothyroid and cricoarytenoid joints. These cartilages support the soft tissues of the laryngeal cavity, thereby protecting this vital passageway for unencumbered movement of air into and out of the lower airway. The thyroid cartilage is composed of two quadrangular plates that are united at midline in an angle called the thyroid angle or laryngeal prominence. In the male, the junction of the laminae forms an acute angle, while in the female it is obtuse. This sexual dimorphism emerges after puberty. The coid cartilage is signet ring shaped and sits on top of the first ring of the trachea, ensuring continuity of the airway from the larynx into the trachea (the origin of the lower respiratory tract). The epiglottis is a flexible leaf-shaped cartilage whose deformability results from its elastic cartilage composition. During swallowing, the epiglottis closes over the entrance into the laryngeal cavity, thus preventing food and liquids from passing into the laryngeal cavity, which could obstruct the airway and interfere with breathing. The arytenoid cartilages are interconnected to the coid cartilage via the cricoarytenoid joint. These pyramid-shaped cartilages serve as points of attachment for the vocal folds, all but one pair of the intrinsic laryngeal muscles, and the vestibular folds.

The thyroid, coid and arytenoid cartilages are interconnected to each other by two movable joints, the cricothyroid and cricoarytenoid joints. The cricothyroid joint joins the thyroid and coid cartilages and allows the coid cartilage to rotate upward toward the coid (Stone and Nuttal, 1974). Since the vocal folds are attached anteriorly to the inside face of the thyroid cartilage and posteriorly to the arytenoid cartilages, which in turn are attached to the upper rim of the coid, this rotation effects lengthening and shortening of the vocal folds, with concomitant changes in tension. Such changes in tension are the principal method of changing the rate of vibration of the vocal folds. The cricoarytenoid joint joins the arytenoid cartilages to the superolateral rim of the coid. Rocking motions of the arytenoids on the upper rim of the coid cartilage allow
the arytenoids and the attached vocal folds to be drawn away (abducted) from midline and brought toward (adducted) midline. The importance of these actions has been emphasized by von Leden and Moore (1961), as they are necessary for developing the transglottal impedances to airflow that are needed to initiate vocal fold vibration. The effect of such movements is to change the size and shape of the glottis, the space between the vocal folds, which is of importance in laryngeal articulation, producing devoicing and pauses, and facilitating modes of vocal attack.

**Laryngeal Cavity.** The laryngeal cartilages surround an irregularly shaped tube called the laryngeal cavity, which forms the interior of the larynx (Fig. 2). It extends from the laryngeal inlet (laryngeal aditus), through which it communicates with the hypopharynx, to the level of the inferior border of the cricoid cartilage. Here the laryngeal cavity is continuous with the lumen of the trachea. The walls of the laryngeal cavity are formed by fibroelastic tissues lined with epithelium. These fibroelastic tissues (quadrangular membrane and conus elasticus) restore the dimensions of the laryngeal cavity, which become altered through muscle activity, passive stretch from adjacent structures, and aeromechanical forces.

![Figure 2. The laryngeal cavity, as viewed posteriorly. (From Kahane, J. C. [1988]. Anatomy and physiology of the organs of the peripheral speech mechanism. In N. J. Lass, L. L. McReynolds, J. L. Northern, and D. E. Yoder [Eds.], Handbook of speech-language pathology and audiology. Toronto: B. C. Decker. Reproduced with permission.)](image-url)

The laryngeal cavity is conventionally divided into three regions. The upper portion is a somewhat expanded supraglottal cavity or vestibule whose walls are reinforced by the quadrangular membrane. The middle region, called the glottal region, is bounded by the vocal folds; it is the narrowest portion. The lowest region, the infraglottal or subglottal region, is bounded by the conus elasticus. The area of primary laryngeal valving is the glottal region, where the shape and size of the *rima glottidis* or *glottis* (space between the vocal folds) is modified during respiration, vocalization, and sphincteric closure. The rima glottidis consists of an intramembranous portion, which is bordered by the soft tissues of the vocal folds, and an intracartilaginous portion, the posterior two-fifths of the rima glottidis, which is located between the vocal processes and the bases of the arytenoid cartilages. The anterior two-thirds of the glottis is an area of dynamic change occasioned by the positioning and aerodynamic displacement of the vocal folds. The overall dimensions of the intracartilaginous glottis remain relatively stable except during strenuous sphincteric valving.

The epithelium that lines the laryngeal cavity exhibits regional specializations. Stratified squamous epithelium covers surfaces subjected to contact, compressive, and vibratory forces. Typical respiratory epithelium (pseudostratified ciliated columnar epithelium with goblet cells) is plentiful in the laryngeal cavity and lines the supraglottis, ventricles, and nonvibrating portions of the vocal folds; it also provides filtration and moisturization of flowing air. The epithelium and immediately underlying connective tissue form the muscosa, which is supplied by an array of sensory receptors sensitive to pressure, chemical, and tactile stimuli, pain, and direction and velocity of airflow (Wyke and Kirchner, 1976). These receptors are innervated by sensory branches from the superior and recurrent laryngeal nerves. They are essential components of the exquisitely sensitive protective reflex mechanism within the larynx that includes initiating coughing, throat clearing, and sphincteric closure.

**Laryngeal Muscles.** The larynx is acted upon by extrinsic and intrinsic laryngeal muscles (Tables 1 and 2). The *extrinsic laryngeal muscles* are attached at one end to the larynx and have one or more sites of attachment to a distant site (e.g., the sternum or hyoid bone). The suprahyoid and infrahyoid muscles attach to the hyoid bone and are generally considered extrinsic laryngeal muscles (Fig. 3). Although these muscles do not attach to the larynx, they influence laryngeal position in the neck through their action on the hyoid bone. The thyroid cartilage is connected to the hyoid bone by the thyrohyoid membrane and ligaments. The larynx is moved through displacement of the hyoid bone. The suprahyoid and infrahyoid muscles also stabilize the hyoid bone, allowing other muscles in the neck to act directly on the laryngeal cartilages. The suprahyoid and infrahyoid muscles are innervated by a combination of cranial and spinal nerves. Cranial nerves V and VII
Table 1. Morphological Characteristics of the Suprahyoid and Infrahyoid Muscles

<table>
<thead>
<tr>
<th>Muscles</th>
<th>Origin</th>
<th>Insertion</th>
<th>Function</th>
<th>Innervation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Suprahyoid Muscles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior digastric</td>
<td>Digastric fossa of mandible</td>
<td>Body of hyoid bone</td>
<td>Raises hyoid bone</td>
<td>Cranial nerve V</td>
</tr>
<tr>
<td>Posterior digastric</td>
<td>Mastoid notch of temporal bone</td>
<td>To hyoid bone via an intermediate tendon</td>
<td>Raises and retracts hyoid bone</td>
<td>Cranial nerve VII</td>
</tr>
<tr>
<td>Stylohyoid</td>
<td>Posterior border of styloid process</td>
<td>Body of hyoid</td>
<td>Raises hyoid bone</td>
<td>Cranial nerve VII</td>
</tr>
<tr>
<td>Mylohyoid</td>
<td>Mylohyoid line of mandible</td>
<td>Median raphe, extending from deep surface of</td>
<td>Raises hyoid bone</td>
<td>Cranial nerve V</td>
</tr>
<tr>
<td>Geniohyoid</td>
<td>Inferior pair of genial tubercles of mandible</td>
<td>Anterior surface of body of hyoid bone</td>
<td>Raises hyoid bone and draws it forward</td>
<td>Cervical nerve I carried</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>via descendens hypoglossi</td>
</tr>
<tr>
<td><strong>Infrahyoid Muscles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sternohyoid</td>
<td>Deep surface of manubrium; medial end of clavicle</td>
<td>Medial portion of inferior surface of body of hyoid bone</td>
<td>Depresses hyoid bone</td>
<td>Ansa cervicalis</td>
</tr>
<tr>
<td>Omohyoid</td>
<td>From upper border of scapula (inferior belly) into tendon issuing superior belly</td>
<td>Inferior aspect of body of hyoid bone</td>
<td>Depresses hyoid bone</td>
<td>Cervical nerves I–III carried by the ansa cervicalis</td>
</tr>
<tr>
<td>Sternothyroid</td>
<td>Posterior surface of manubrium; edge of first costal cartilage</td>
<td>Oblique line of thyroid cartilage</td>
<td>Lowers hyoid bone; stabilizes hyoid bone</td>
<td>Ansa cervicalis</td>
</tr>
<tr>
<td>Thyrohyoid</td>
<td>Oblique line of thyroid cartilage</td>
<td>Lower border of body and greater wing of hyoid bone</td>
<td>When larynx is stabilized, lowers hyoid bone; when hyoid is fixed, larynx is raised</td>
<td>Cervical nerve I, through descendens hypoglossi</td>
</tr>
</tbody>
</table>

Table 2. Morphological Characteristics of the Intrinsic Laryngeal Muscles

<table>
<thead>
<tr>
<th>Muscle</th>
<th>Origin</th>
<th>Insertion</th>
<th>Function</th>
<th>Innervation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cricothyroid</td>
<td>Lateral surface of cricoid cartilage arch; fibers divide into upper portion (pars recta) and lower portion (pars obliqua)</td>
<td>Pars recta fibers attach to anterior lateral half of inferior border of thyroid cartilage; pars obliqua fibers attach to anterior margin of inferior corner of thyroid cartilage</td>
<td>Rotational approximation of the cricoid and thyroid cartilages; lengthens and tenses vocal folds</td>
<td>External branch of superior laryngeal nerve (cranial nerve X)</td>
</tr>
<tr>
<td>Lateral cricoarytenoid</td>
<td>Upper border of arch of cricoid cartilage</td>
<td>Anterior aspect of muscular process of arytenoid cartilage</td>
<td>Adducts vocal folds; closes rima glottidis</td>
<td>Recurrent laryngeal nerve (cranial nerve X)</td>
</tr>
<tr>
<td>Posterior cricoarytenoid</td>
<td>Cricoid lamina</td>
<td>Muscular process of arytenoid cartilage</td>
<td>Abducts vocal folds; opens rima glottidis</td>
<td>Recurrent laryngeal nerve (cranial nerve X)</td>
</tr>
<tr>
<td>Interarytenoid Transverse fibers</td>
<td>Horizontally coursing fibers extending between the dorso-lateral ridges of each arytenoid cartilage</td>
<td>Dorsolateral ridge of opposite arytenoid cartilage</td>
<td>Approximates bases of arytenoid cartilages, assists vocal fold adduction</td>
<td>Recurrent laryngeal nerve (cranial nerve X)</td>
</tr>
<tr>
<td>Oblique fibers</td>
<td>Obliquely coursing fibers from base of one arytenoid cartilage</td>
<td>Inserts onto apex of opposite arytenoid cartilage</td>
<td>Same as transverse fibers</td>
<td>Recurrent laryngeal nerve (cranial nerve X)</td>
</tr>
<tr>
<td>Thyroarytenoid</td>
<td>Deep surface of thyroid cartilage at midline</td>
<td>Fovea oblonga of arytenoid cartilage; vocalis fibers attach close to vocal process; muscularis fibers attach more laterally</td>
<td>Adduction, tensor, relaxer of vocal folds (depending on what parts of muscles are active)</td>
<td>Recurrent laryngeal nerve (cranial nerve X)</td>
</tr>
</tbody>
</table>
Among the most important functional and biomechanical outcomes of the actions of the intrinsic laryngeal muscles are (1) abduction and adduction of the vocal folds, (2) changing the position of the laryngeal cartilages relative to each other, (3) transiently changing the dimensions and physical properties of the vocal folds (i.e., length, tension, mass per unit area, compliance, and elasticity), and (4) modifying laryngeal airway resistance by changing the size or shape of the glottis.

The intrinsic laryngeal muscles are innervated by nerve fibers carried in the trunk of the vagus nerve. These branches are usually referred to as the superior and inferior laryngeal nerves. The cricothyroid muscle is innervated by the superior laryngeal nerve, while all other intrinsic laryngeal muscles are innervated by the inferior (recurrent) laryngeal nerve. Sensory fibers from these nerves supply the entire laryngeal cavity.

Histochemical studies of intrinsic laryngeal muscles (Matzelt and Vosteen, 1963; Rosenfield et al., 1982) have enabled us to appreciate the unique properties of the intrinsic muscles. The intrinsic laryngeal muscles contain, in varying proportions, fibers that control fine movements for prolonged periods (type 1 fibers) and fibers that develop tension rapidly within a muscle (type 2 fibers). In particular, laryngeal muscles differ from the standard morphological reference for striated muscles, the limb muscles, in several ways: (1) they typically have a smaller mean diameter of muscle fibers; (2) they are less regular in shape; (3) the muscle fibers are generally uniform in diameter across the various intrinsic muscles; (4) individual muscle fibers tend not to be uniform in their directionality within a fascicle but exhibit greater variability in the course of muscle fibers, owing to the tendency for fibers to intermingle in their longitudinal and transverse planes; and (5) laryngeal muscles have a greater investment of connective tissues.

**Vocal Folds.** The vocal folds are multilayered vibrators, not a single homogeneous band. Hirano (1974) showed that the vocal folds are composed of several layers of tissues, each with different physical properties and only 1.2 mm thick. The vocal fold consists of one layer of epithelium, three layers of connective tissue (lamina propria), and the vocalis fibers of the thyroarytenoid muscle (Fig. 5). Based on examination of ultra-high-speed films and biomechanical testing of the vocal folds, Hirano (1974) found that functionally, the epithelium and superficial layer of the lamina propria form the cover, which is the most mobile portion of the vocal fold. Wave-like mucosal disturbances travel along the surface during sound production. These movements are essential for developing the agitation and patterning of air molecules in transglottal airflow during voice production. The superficial layer of the lamina propria is composed of sparse amounts of loosely interwoven collagenous and elastic fibers. This area, also known as Reinke’s space, is important clinically because it is the principal site of swelling or edema formation in the vocal folds following vocal abuse or in laryngitis. The intermediate and deep layers of the lamina propria are called the transition. The vocal ligament is formed from elastic and collagogenous fibers in these layers. It provides...
Figure 4. The intrinsic laryngeal muscles as shown in lateral (A), posterior (B), and superior (C) views. (From Kahane, J. C. [1988]. Anatomy and physiology of the organs of the peripheral speech mechanism. In N. J. Lass, L. L. McReynolds, J. L. Northern, and D. E. Yoder [Eds.], Handbook of speech-language pathology and audiology. Toronto: B. C. Decker. Reproduced with permission.)
resiliency and longitudinal stability to the vocal folds during voice production. The transition is stiffer than the cover but more pliant than the vocalis muscle fibers, which form the body of the vocal folds. These muscle fibers are active in regulating fundamental frequency by influencing the tension in the vocal fold and the compliance and elasticity of the vibrating surface (cover).

See also voice production: physics and physiology.

—Joel C. Kahane

References


Further Readings


Assessment of Functional Impact of Voice Disorders

Introduction

Voice disorders occur in approximately 6% of all adults and in as many as 12% of children. Within the adult group, specific professions report the presence of a voice problem that interferes with their employment. As many as 50% of teachers and 33% of secretaries complain of voice problems that restrict their ability to work or to function in a normal social environment (Smith et al., 1998). The restriction of work, or lifestyle, due to a voice disorder has gone virtually undocumented until recently. While voice scientists and clinicians have focused most of their energy, talent, and time on diagnosing and measuring the severity of voice disorders with various perceptual, acoustic, or physiological instruments, little attention has been given to the effects of a voice disorder on the daily needs of the patient. Over the past few years, interest has increased in determining the functional impact of the voice disorder due to the Internet in using patient-based outcome measures to establish efficacy of treatments and the desire to match treatment needs with patient’s needs. This article reviews the evolution of the assessment of functional impact of voice disorders and selected applications of those assessments. Assessment of the physiological consequences of voice disorders has evolved from a strong interest in the relationship of communication ability to global quality-of-life measurement. Hassan and Weymuller (1993), List et al. (1998), Picarillo (1994), and Murry et al. (1998) have all demonstrated that voice communication is an essential element in patients’ perception of their quality of life following treatment for head and neck cancer. Patient-based assessment of voice handicap has been lacking in the area of noncancerous voice disorders. The developments and improvements of software for assessing acoustic objective measures of voice and relating measures of abnormal voices to normal voices have gone on for a number of years. However, objective measures primarily assess specific treatments and do not encompass functional outcomes from the patient’s perspective. These measures do not necessarily discriminate the severity of handicap as it relates to specific professions. Objective test batteries are useful to quantify disease severity (Rosen, Lombard, and Murry, 2000), categorize acoustic/physiological profiles of the disease (Hartl et al., 2001), and measure changes that occur as a result of treatment (Dejonker, 2000). A few objective and subjective measures are correlated with the diagnosis of the voice disorder (Wolfe, Fitch, and Martin, 1997), but until recently, none have been related to the patient’s perception of the severity of his or her problem. This latter issue is important in all diseases and disorders when life is not threatened since it is ultimately the patient’s perception of disease severity and his or her motivation to seek treatment that dictates the degree of treatment success.

Functional impact relates to the degree of handicap or disability. Accordingly, there are three levels of a disorder: impairment, disability, and handicap (World Health Organization, 1980). Handicap is the impact of the impairment of the disability on the social, environmental, or economic functioning of the individual. Treatment usually relates to the physical well-being of a patient, and it is this physical well-being that generally takes priority when attempting to assess the severity of the handicap. A more comprehensive approach might seek to address the patient’s own impression of the severity of the disorder and how the disorder interferes with the individual’s professional and personal lifestyle.

Measurement of functional impact is somewhat different from assessment of disease status in that it does not directly address treatment efficacy, but rather addresses the value of a particular treatment for a particular individual. This may be considered treatment effectiveness. Efficacy, on the other hand, looks at whether or not a treatment can produce an expected result based on previous studies. Functional impact relates to the degree of impact a disorder has on an individual patient, not necessarily to the severity of the disease.

Voice Disorders and Outcomes Research

Assessment of functional impact on the voice is barely beyond the infancy stage. Interest in the issues relating to functional use of the voice stems from the development of instruments to measure all aspects of vocal function related to the patient, the disease, and the treatment.
Moreover, there are certain parameters of voice disorders that cannot be easily measured in the voice laboratory, such as endurance, acceptance of a new voice, and vocal effectiveness.

The measurement of voice handicap must take into account issues such as “can the person teach in the classroom all day?” or “can a shop foreman talk loud enough to be heard over the noise of factory machines?” An outcome measure that takes into account the patient’s ability to speak in the classroom or a factory will undoubtedly provide a more accurate assessment of voice handicap (although not necessarily an accurate assessment of the disease, recovery from disease, or quality of voice) than the acoustic measures obtained in the voice laboratory. Thus, patient-based measures of voice handicap provide significant information that cannot be obtained from biological and physiologic variables traditionally used in voice assessment models.

Voice handicap measures may measure an individual’s perceived level of general health, an individual’s quality of life, her ability to continue with her current employment versus opting for a change in employment, her satisfaction with treatment regardless of the disease state, or the cost of the treatment. Outcome of treatment for laryngeal cancer is typically measured using Kaplan-Meyer curves (Adelstein et al., 1990). While this tool measures the disease-related status of the patient, it does not presume to assess overall patient satisfaction with treatment. Rather, the degree to which swallowing status improves and voice communication returns to normal are measured by instruments that generally focus on quality of life (McHorney et al., 1993).

Voice disorders are somewhat different than the treatment of a life threatening disease such as laryngeal cancer. Treatment that involves surgery, pharmacology, or voice therapy requires the patient’s full cooperation throughout the course of treatment. The quality and accuracy of surgery or the level of voice therapy may not necessarily reflect the long-term outcome if the patient does not cooperate with the treatment procedure. Assessment of voice handicap involves the patient’s ability to use his or her voice under normal circumstances of social and work-related speaking situations. The voice handicap will be reflected to the extent that the voice is usable in those situations.

Outcome Measures: General Health Versus Specific Disease

There are two primary ways to assess the handicap of a voice disorder. One is to look at the patient’s overall well-being. The other is to compare his or her voice to normal voice measures. The first usually encompasses social factors as well as physical factors that are related to the specific disorder. One measure that has been used to look at the effect of disease on life is the Medical Outcomes Study (MOS), a 36-item short-form general health survey (McHorney et al., 1993). The 36-item short form, otherwise known as SF-36, measures eight areas of health that are commonly affected or changed by diseases and treatments: physical functioning, role functioning, bodily pain, general health, vitality, social functioning, mental health, and health transition. The SF-36 has been used for a wide range of disease-specific topics once it was shown to be a valid measure of the degree of general health. The SF-36 is a pencil-and-paper test that has been used in numerous studies for assessing outcomes of treatment. In addition, because each scale has been determined to be a reliable and valid measure of health in and of itself, this assessment has been used to validate other assessments of quality of life and handicap that are disease specific. However, one of the difficulties with using such a test for a specific disease is that one or more of the subscales may not be important or appropriate. For example, when considering certain voice disorders, the subscale of the SF-36 known as bodily pain may not be quite appropriate. Thus, the SF-36 is not a direct assessment of voice handicap but rather a general measure of well-being.

The challenge to develop a specific scale related to a specific organ function such as a scale for voice disorders presents problems unlike the development of the SF-36 or other general quality-of-life scales.

Assessing Voice Handicap

Currently there are no federal regulations defining voice handicap, unlike the handicap measures associated with hearing loss, which is regulated by the Department of Labor. The task of measuring the severity of a voice disorder may be somewhat difficult because of the areas that are affected, namely emotional, physical, functional, economic, etc. Moreover, as already indicated, while measures such as a perceptual judgments of voice characteristics, videostroboscopic visual perceptual findings, acoustic perceptual judgments, as well as physiological measures objectively obtained provide some input as to the severity of the voice compared to normal, these measures do not provide insight as to the degree of handicap and disability that a specific patient is experiencing. It should be noted, however, that there are handicap/disability measures developed for other aspects of communication, namely hearing loss and dizziness (Newman et al., 1990; Jacobson et al., 1994). These measures have been used to quantify functional outcome following various interventions in auditory function.

Development of the Voice Handicap Index

In 1997, Jacobson and her colleagues proposed a measure of voice handicap known as the Voice Handicap Index (VHI) (Jacobson et al., 1998). This patient self-assessment tool consists of ten items in each of three domains: emotional, physical, and functional aspects of voice disorders. The functional subscale includes statements that describe the impact of a person’s voice on his daily activities. The emotional subscale indicates the patient’s affective responses to the voice disorder. The items in the physical subscale are statements that relate to either the patient’s perception of laryngeal discomfort or the voice output characteristics such as too low or too high a pitch. From an original 85-item list, a 30-item
questionnaire using a five-point response scale from 0, indicating he “never” felt this about his voice problem to 4, where he “always” felt this to be the case, was finally obtained. This 30-item questionnaire was then assessed for test-retest stability in total as well as the three subscales, and was validated against the SF-36. A shift in the total score of 18 points or greater is required in order to be certain that a change is due to intervention and not to unexplained variability. The Voice Handicap Index was designed to assess all types of voice disorders, even those encountered by tracheoesophageal speakers. A detailed analysis of patient data using this test has recently been published (Benninger et al., 1998).

Since the VHI has been published, others have proposed similar tests of handicap. Hogikian (1999) and Glicklich (1999) have both demonstrated their assessment tools to have validity and reliability in assessing a patient’s perception of the severity of a voice problem.

One of the additional uses of the VHI as suggested by Benninger and others is to assess measures after treatment (1998). Murry and Rosen (2001) evaluated the VHI in three groups of speakers to determine the relative severity of voice disorders in patients with muscular tension dysphonia (MTD), benign vocal fold lesions (polyps/cysts), and vocal fold paralysis prior to and following treatments. Figure 1 shows that subjects with vocal fold paralysis displayed the highest self-perception of handicap both before and after treatment. Subjects with benign vocal fold lesions demonstrated the lowest perception of handicap severity before and after treatment. It can be seen that in general, there was a 50% or greater improvement in the mean VHI for the combined groups. However, the patients with vocal fold paralysis initially began with the highest pretreatment VHI and remained with the highest VHI after treatment. Although the VHI scores following treatment were significantly lower, there still remained a measure of handicap in all subjects. Overall, in 81% of the patients, there was a perception of significantly reduced voice handicap, either because of surgery, voice therapy, or a combination of both.

The same investigators examined the application of the VHI to a specific group of patients with voice disorders, singers (Murry and Rosen, 2000). Singers are unique in that they often complain of problems related only to their singing voice. Murry and Rosen examined 73 professional and 33 nonprofessional singers and compared them with a control group of 369 nonsingers.

The mean VHI score for the 106 singers was 34.7, compared with a mean of 53.2 for the 336 nonsingers. The VHI significantly separated singers from nonsingers in terms of severity. Moreover, the mean VHI score for the professional singers was significantly lower (31.0 vs. 43.2) than for the recreational singers. Although lower VHI scores were found in singers than in nonsingers, this does not imply that the VHI is not a useful instrument for assessing voice problems in singers. On the contrary, several questions were singled out as specifically sensitive to singers. The findings of this study should alert clinicians that the use of the VHI points to the specific needs as well as the seriousness of a singer’s handicap.

The quality of voice may be mildly disordered, the voice handicap may be significant.

Recently, Rosen and Murry (in press) presented reliability data on a revised 10-question VHI. The results suggest that a 10-question VHI produces a highly correlated with the original VHI. The 10-item questionnaire provides a quick, reliable assessment of the patient’s perception of voice handicap.

Other measures of voice outcome have been proposed and studied. Recently, Gliklich, Glovsky, and Montgomery examined outcomes in patients with vocal fold paralysis (Hogikyan and Sethuraman, 1999). The instrument, which contains five questions, is known as the Voice Outcome Survey (VOS). Overall reliability of the VOS was related to the subscales of the SF-36 for a group of patients with unilateral vocal fold paralysis.

Additional work has been done by Hogikyan (1999). These authors presented a measure of voice-related quality of life (VR-QOL). They also found that this self-administered 10-question patient assessment of severity was related to changes in treatment. Their subjects consisted primarily of unilateral vocal fold paralysis patients and showed a significant change from pre- to post-treatment.

A recent addition to functional assessment is the Voice Activity and Participation Profile (VAPP). This tool assesses the effects voice disorders have on limiting and participating in activities which require use of the voice (Ma and Yiu, 2001). Activity limitation refers to constraints imposed on voice activities and participation restriction refers to a reduction or avoidance of voice activities. This 28-item tool examines five areas: self-perceived severity of the voice problem; effect on the job; effect on daily communication; effect on social communication; and effect on emotion. The VAPP has been found to be a reliable and valid assessment tool for assessing self-perceived voice severity as it relates to activity and participation in vocal activities.

---

**Figure 1.** Pre- and post-treatment voice handicap scores for selected populations.
**Summary**

The study of functional voice assessment to identify the degree of handicap is novel for benign voice disorders. For many years, investigators have focused on acoustic and aerodynamic measures of voice production to assess change in voice following treatment. These measures, although extremely useful in understanding treatment efficacy, have not shed significant light on patients’ perception of their disorder. Measures such as the VHI, VOS, and VR-QOL have demonstrated that regardless of age, sex, or disease type, the degree of handicap can be identified. Furthermore, treatment for these handicaps can also be assessed in terms of effectiveness for the patient. Patients’ self-assessment of perceived severity also allows investigators to make valid comparisons of the impact of an intervention for patients who use their voices in different environments and the patients’ perception of the treatment from a functional perspective. Assessment of voice based on a patient’s perceived severity and the need to recover vocal function may be the most appropriate manner to assess severity of voice handicap.

—Thomas Murry and Clark A. Rosen

**References**


**Electroglottographic Assessment of Voice**

A number of instruments can be used to help characterize the behavior of the glottis and vocal folds during phonation. The signals derived from these instruments are called *glottographic waveforms* or *glottograms* (Titze and Talkin, 1981). Among the more common glottograms are those that track change in glottal flow, via inverse filtering; glottal width, via kymography; glottal area, via photoglottography; and vocal fold movement, via ultrasonography (Baken and Orlikoff, 2000). Such signals can be used to obtain several different physiological measures, including the glottal open quotient and the maximum flow declination rate, both of which are highly valuable in the assessment of vocal function. Unfortunately, the routine application of these techniques has been hampered by the cumbersome and time-consuming way in which these signals must be acquired, conditioned, and analyzed. One glottographic method,
**Electroglottography (EGG)** has emerged as the most commonly used technique, for several reasons: (1) it is noninvasive, requiring no probe placement within the vocal tract; (2) it is easy to acquire, alone or in conjunction with other speech signals; and (3) it offers unique information about the mucoidulatory behavior of the vocal folds, which contemporary theory suggests is a critical element in the assessment of voice production.

Electroglottography (known as electrolaryngography in the United Kingdom) is a plethysmographic technique that entails fixing a pair of surface electrodes to each side of the neck at the thyroid lamina, approximating the level of the vocal folds. An imperceptible low-amplitude, high-frequency current is then passed between these electrodes. Because of their electrolyte content, tissue and body fluids are relatively good conductors of electricity, whereas air is a particularly poor conductor. When the vocal folds separate, the current path is forced to circunvent the glottal air space, decreasing effective voltage. Contact between the vocal folds affords a conduit through which current can take a more direct route across the neck. Electrical impedance is thus highest when the current path must completely bypass an open glottis and progressively decreases as greater contact between the vocal folds is achieved. In this way, the voltage across the neck is modulated by the contact of the vocal folds, forming the basis of the EGG signal. The glottal region, however, is quite small compared with the total region through which the current is flowing. In fact, most of the changes in transcervical impedance are due to strap muscle activity, laryngeal height variation induced by respiration and articulation, and pulsatile blood volume changes. Because increasing and decreasing vocal fold contact has a relatively small effect on the overall impedance, the electroglottogram is both high-pass filtered to remove the far slower nonphonatory impedance changes and amplified to boost the laryngeal contribution to the signal. The result is a waveform—sometimes designated $L_x$—that varies chiefly as a function of vocal fold contact area (Gilbert, Potter, and Hoodin, 1984).

First proposed by Fabre in 1957 as a means to assess laryngeal physiology, the clinical potential of EGG was recognized by the mid-1960s. Interest in EGG increased in the 1970s as the importance of mucosal wave dynamics for vocal fold vibration was confirmed, and accelerated greatly in the 1980s with the advent of personal computers and commercially available EGGs that were technologically superior to previous instruments. Today, EGG has a worldwide reputation as a useful tool to supplement the evaluation and treatment of vocal pathology. The clinical challenge, however, is that a valid and reliable EGG assessment demands a firm understanding of normal vocal fold vibratory behavior along with recognition of the specific capabilities and limitations of the technique.

Instead of a simple mediolateral oscillation, the vocal folds engage in a quite complex undulatory movement during phonation, such that their inferior margins approximate before the more superior margins make contact. Because EGG tracks effective medial contact area, the pattern of vocal fold vibration can be characterized quite well (Fig. 1). The contact pattern will vary as a consequence of several factors, including bilateral vocal fold mass and tension, medial compression, and the anatomy and orientation of the medial surfaces. Considerable research has been devoted to establishing the important features of the EGG and how they relate to specific aspects of vocal fold status and behavior. Despite these efforts, however, the contact area function is far from perfectly understood, especially in the face of pathology. Given the complexity of the “rolling and peeling” motion of the glottal margins and the myriad possibilities for abnormality of tissue structure or biomechanics, it is not surprising that efforts to formulate simple rules relating abnormal details to specific pathologies have not met with notable success. In short, the clinical value of EGG rests in documenting the vibratory consequence of pathology rather than in diagnosing the pathology itself.
Using multiple glottographic techniques, Baer, Lofqvist, and McGarr (1983) demonstrated that, for normal modal-register phonation, the “depth of closure” was very shallow just before glottal opening and quite deep soon after closure was initiated. Most important, they showed that the instant at which the glottis first appears occurs sometime before all contact is lost, and that the instant of glottal closure occurs sometime after the vocal folds first make contact. Thus, although the EGG is sensitive to the depth of contact, it cannot be used to determine the width, area, or shape of the glottis. For this reason, EGG is not a valid technique for the measurement of glottal open time or, therefore, the open quotient. Likewise, since EGG does not specify which parts of the vocal folds are in contact, it cannot be used to measure glottal closed time, nor can it, without additional evidence, be used to determine whether maximal vocal fold contact indeed represents complete obliteration of the glottal space. Identifying the exact moment when (and if) all medial contact is lost has also proved particularly problematic. Once the vocal folds do lose contact, however, it can no longer be assumed that the EGG signal conveys any information whatsoever about laryngeal behavior. During such intervals, the signal may vary solely as a function of the instrument’s automatic gain control and filtering (Rothenberg, 1981).

Although the EGG provides useful information only about those parts of the vibratory cycle during which there is some vocal fold contact, these characteristics may provide important clinical insight, especially when paired with videostroboscopy and other data traces. EGG, with its ability to demonstrate contact change in both the horizontal and vertical planes, can quite effectively document the normal voice registers (Fig. 2) as well as abnormal and unstable modes of vibration (Fig. 3). However, to qualitatively assess EGG wave characteristics and to derive useful indices of vocal fold contact behavior, it may be best to view the EGG in terms of a vibratory cycle composed of a contact phase and a minimal-contact phase (see Fig. 1). The contact phase includes intervals of increasing and decreasing contact, whereas the peak represents maximal vocal fold contact, and, presumably, maximal glottal closure. The minimal-contact phase is that portion of the EGG wave during which the vocal folds are probably not in contact. Much clinical misinterpretation can be avoided if no attempt is made to equate the vibratory contact phase with the glottal closed phase or the minimal-contact phase with the glottal open phase.

For the typical modal-register EGG, the contact phase is asymmetrical; that is, the increase in contact takes less time than the interval of decreasing contact. The degree of contact asymmetry is thought to vary not only as a consequence of vocal fold tension but also as a function of vertical mucosal convergence and dynamics (i.e., phasing; Titze, 1990). A dimensionless ratio, the contact index (CI), can be used to assess contact symmetry (Orlikoff, 1991). Defined as the difference between the increasing and decreasing contact durations divided by the duration of the contact phase, CI will vary between –1 for a contact phase maximally skewed to the left and +1 for a contact phase maximally skewed to the right. For normal modal-register phonation, CI varies between –0.6 and –0.4 for both men and women, but, as can be seen in Figure 2, it is markedly different for other voice registers. Pulse-register EGGs typically have CIs in the vicinity of –0.8, whereas in falsetto it would not be uncommon to have a CI that approximates zero, indicating a symmetrical or nearly symmetrical contact phase.

Another EGG measure that is gaining some currency in the clinical literature is the contact quotient (CQ). Defined as the duration of the contact phase relative to the period of the entire vibratory cycle, there is evidence from both in vivo testing and mathematical modeling to suggest that CQ varies with the degree of medial compression of the vocal folds (see Fig. 3) along a hypoadducted “loose” (or “breathy”) to a hyperadducted “tight” (or “pressed”) phonatory continuum (Rothenberg and Mahshie, 1988; Titze, 1990). Under typical vocal circumstances, CQ is within the range of 40%–60%, and despite the propensity for a posterior glottal
chink in women, there does not seem to be a significant sex effect. This is probably due to the fact that EGG (and thus the CQ) is insensitive to glottal gaps that are not time varying. Unlike men, however, women tend to show an increase in CQ with vocal F0. It has been conjectured that this may be the result of greater medial compression employed by women at higher F0s that serves to diminish the posterior glottal gap. Nonetheless, a strong relationship between CQ and vocal intensity has been documented in both men and women, consistent with the known relationship between vocal power and the adductory presetting of the vocal folds. Because vocal intensity is also related to the rate of vocal fold contact (Kakita, 1988), there have been some preliminary attempts to derive useful EGG measures of the contact rise time.

Because EGG is relatively unaffected by vocal tract resonance and turbulence noise (Orlikoff, 1995), it allows evaluation of vocal fold behavior under conditions not well-suited to other voice assessment techniques. For this reason, and because the EGG waveshape is a relatively simple one, the EGG has found some success both as a trigger signal for laryngeal videostroboscopy and as a means to define and describe phonatory onset, offset, intonation, voicing, and fluency characteristics. In fact, EGG has, for many, become the preferred means by which to measure vocal fundamental frequency and jitter.

In summary, EGG provides an innocuous, straightforward, and convenient way to assess vocal fold vibration through its ability to track the relative area of contact. Although it does not supply valid information about the opening and closing of the glottis, the technique affords a unique perspective on vocal fold behavior. When conservatively interpreted, and when combined with other tools of laryngeal evaluation, EGG can substantially further the clinician’s understanding of the malfunctioning larynx and play an effective role in therapeutics as well.

See also ACOUSTIC ASSESSMENT OF VOICE.

—Robert F. Orlikoff

References


Further Readings


Functional Voice Disorders

The human voice is acutely responsive to changes in emotional state, and the larynx plays a prominent role as an instrument for the expression of intense emotions such as fear, anger, grief, and joy. Consequently, many regard the voice as a sensitive barometer of emotions and the larynx as the control valve that regulates the release of these emotions (Aronson, 1990). Furthermore, the voice is one of the most individual and characteristic expressions of a person—a “mirror of personality.” Thus, when the voice becomes disordered, it is not uncommon for clinicians to suggest personality traits, psychological factors, or emotional or inhibitory processes as primary causal mechanisms. This is especially true in the case of functional dysphonia or aphonia, in which no visible structural or neurological laryngeal pathology exists to explain the partial or complete loss of voice.

Functional dysphonia, which may account for more than 10% of cases referred to multidisciplinary voice clinics, occurs predominantly in women, commonly follows upper respiratory infection symptoms, and varies in its response to treatment (Bridger and Epstein, 1983; Schalen and Andersson, 1992). The term functional implies a voice disturbance of physiological function rather than anatomical structure. In clinical circles, functional is usually contrasted with organic and often carries the added meaning of psychogenic. Stress, emotional, and psychological conflict are frequently presumed to cause or exacerbate functional symptoms.

Some confusion surrounds the diagnostic category of functional dysphonia because it includes an array of medically unexplained voice disorders: psychogenic, conversion, hysterical, tension-fatigue syndrome, hyperkinetic, muscle misuse, and muscle tension dysphonia. Although each diagnostic label implies some degree of etiologic heterogeneity, whether these disorders are qualitatively different and etiologically distinct remains unclear. When applied clinically, these various labels frequently reflect clinician supposition, bias, or preference. Voice disorder taxonomies have yet to be adequately operationalized; consequently, diagnostic categories often lack clear thresholds or discrete boundaries to determine patient inclusion or exclusion. To improve precision, some clinicians prefer the term psychogenic voice disorder, to put the emphasis on the psychological origins of the disorder. According to Aronson (1990), a psychogenic voice disorder is synonymous with a functional one but offers the clinician the advantage of stating confidently, after an exploration of its causes, that the voice disorder is a manifestation of one or more forms of psychological disequilibrium. At the purely phenomenological level there may be little difference between functional and psychogenic voice disorders. Therefore, in this discussion, the terms functional and psychogenic will be used synonymously, which reflects current trends in the clinical literature (nosological imprecision notwithstanding).
In clinical practice, “psychogenic voice disorder” should not be a default diagnosis for a voice problem of undetermined cause. Rather, at least three criteria should be met before such a diagnosis is offered: symptom psychogenicity, symptom incongruity, and symptom reversibility (Sapir, 1995). Symptom psychogenicity refers to the finding that the voice disorder is logically linked in time of onset, course, and severity to an identifiable psychological antecedent, such as a stressful life event or interpersonal conflict. Such information is acquired through a complete case history and psychosocial interview. Symptom incongruity refers to the observation that the vocal symptoms are physiologically incompatible with existing or suspected disease, are internally inconsistent, and are incongruent with other speech and language characteristics. An often cited example of symptom incongruity is complete aphonia (whispered speech) in a patient who has a normal throat clear, cough, laugh, or hum, whereby the presence of such normal nonspeech vocalization is at odds with assumptions regarding neural integrity and function of the laryngeal system. Finally, symptom reversibility refers to complete, sustained amelioration of the voice disorder with short-term voice therapy (usually one or two sessions) or through psychological abreaction. Furthermore, maintaining the voice improvement requires no compensatory effort on the part of the patient. In general, psychogenic dysphonia may be suspected when strong evidence exists for symptom incongruity and symptom psychogenicity, but it is confirmed only when there is unmistakable evidence of symptom reversibility.

A wide array of psychopathological processes contributing to voice symptom formation in functional dysphonia have been proposed. These mechanisms include, but are not limited to, conversion reaction, hysteria, hypochondriasis, anxiety, depression and various personality dispositions or emotional stresses or conflicts that induce laryngeal musculoskeletal tension. Roy and Bless (2000) provide a more complete exploration of the putative psychological and personality processes involved in functional dysphonia, as well as related research.

The dominant psychological explanation for dysphonia unaccounted for by pathological findings is the concept of conversion disorder. According to the DSM-IV, conversion disorder involves unexplained symptoms or deficits affecting voluntary motor or sensory function that suggest a neurological or other general medical condition (American Psychiatric Association, 1994). The conversion symptom represents an unconscious simulation of illness that ostensibly prevents conscious awareness of emotional conflict or stress, thereby displacing the mental conflict and reducing anxiety. When the laryngeal system is involved, the condition is referred to as conversion dysphonia or aphonia. In aphonia, patients lose their voice suddenly and completely and articulate in a whisper. The whisper may be pure, harsh, or sharp, with occasional high-pitched squeaklike traces of phonation. In dysphonia, phonation is preserved but disturbed in quality, pitch, or loudness. Myriad dysphonia types are encountered, including hoarseness (with or without strain), breathiness, and high-pitched falsetto, as well as voice and pitch breaks that vary in consistency and severity.

In conversion voice disorders, psychological factors are judged to be associated with the voice symptoms because conflicts or other stressors precede the onset or exacerbation of the dysphonia. In short, patients convert intrapsychic distress into a voice symptom. The voice loss, whether partial or complete, is also often interpreted to have symbolic meaning. Primary or secondary gains are thought to play an important role in maintaining and reinforcing the conversion disorder. Primary gain refers to anxiety alleviation accomplished by preventing the psychological conflict from entering conscious awareness. Secondary gain refers to the avoidance of an undesirable activity or responsibility and the extra attention or support conferred on the patient.

Butcher and colleagues (Butcher et al., 1987; Butcher, Elias, and Raven, 1993; Butcher, 1995) have argued that there is little research evidence that conversion disorder is the most common cause of functional voice loss. Butcher advised that the conversion label should be reserved for cases of aphonia in which lack of concern and motivation to improve the voice coexists with clear evidence of a temporarily linked psychosocial stressor. In the place of conversion, Butcher (1995) offered two alternative models to account for psychogenic voice loss. Both models minimized the role of primary and secondary gain in maintaining the voice disorder. The first was a slightly reformulated psychoanalytic model that stated, “if predisposed by social and cultural bias as well as early learning experiences, and then exposed to interpersonal difficulties that stimulate internal conflict, particularly in situations involving conflict over self-expression or voicing feelings, intrapsychic conflict or stress becomes channeled into musculoskeletal tension, which physically inhibits voice production” (p. 472). The second model, based on cognitive-behavioral principles, stated that “life stresses and interpersonal problems in an individual predisposed to having difficulties expressing feelings or views would produce involuntary anxiety symptoms and musculoskeletal tension, which would center on and inhibit voice production” (p. 473). Both models clearly emphasized the inhibitory effects of excess laryngeal muscle tension on voice production, although through slightly different causal mechanisms.

Recently, Roy and Bless (2000) proposed a theory that links personality to the development of functional dysphonia. The “trait theory of functional dysphonia” shares Butcher’s (1995) theme of inhibitory laryngeal behavior but attributes this musically inhibited voice production to specific personality types. In brief, the authors speculate that the combination of personality traits such as introversion and neuroticism (trait anxiety) and constraint leads to predictable and conditioned laryngeal inhibitory responses to certain environmental signals or cues. For instance, when undesirable punishing or frustrating outcomes have been paired with previous attempts to speak out, this can lead to musically inhibited voice. The authors contend that this conflict
between laryngeal inhibition and activation (with origins in personality and nervous system functioning) results in elevated laryngeal tension states and can give rise to incomplete or disordered vocalization in a structurally and neurologically intact larynx.

As is apparent from the foregoing discussion, the exquisite sensitivity and prolonged hypercontraction of the intrinsic and extrinsic laryngeal muscles in response to stress, anxiety, depression, and inhibited emotional expression is frequently cited as the common denominator underlying the majority of functional voice problems. Nichol, Morrison, and Rammage (1993) proposed that excess muscle tension arises from overactivity of autonomic and voluntary nervous systems in individuals who are unduly aroused and anxious. They added that such overactivity leads to hypertonicity of the intrinsic and extrinsic laryngeal muscles, resulting in muscle tension dysphonias sometimes associated with adjustment or anxiety disorders, or with certain personality trait disturbances.

Finally, some researchers have noted that their "psychogenic dysphonia and aphonia" patients had an abnormally high number of reported allergy, asthma, or upper respiratory infection symptoms, suggesting a link between psychological factors and respiratory and phonatory disorders (Milutinovic, 1991; Schalen and Andersson, 1992). They have speculated that organic changes in the larynx, pharynx, and nose facilitate the appearance of a functional voice problem; that is, these changes direct the somatization of psychodynamic conflict. Likewise, Rammage, Nichol, and Morrison (1987) proposed that a relatively minor organic change such as edema, infection, or reflux laryngitis may trigger functional misuse, particularly if the individual is exceedingly anxious about his or her voice or health. In a similar vein, the same authors felt that anticipation of poor voice production in hypochondriacal, dependent, or inhibited, stress reactive, socially anxious, nonassertive, and with a tendency toward restraint (Friedl, Friedrich, and Egger, 1990; Gerritsma, 1991; Roy, Bless, and Heisey, 2000a, 2000b).

In conclusion, the larynx can be a site of neuromuscular tension arising from stress, emotional inhibition, fear or threat, communication breakdown, and certain personality types. This tension can produce severely disordered voice in the context of a structurally normal larynx. Although the precise mechanisms underlying and maintaining psychogenic voice problems remain unclear, the voice disorder is a powerful reminder of the intimate relationship between mind and body.

See also Psychogenic Voice Disorders: Direct Therapy.  
—Nelson Roy

References

Hypokinetic Laryngeal Movement Disorders

Hypokinetic laryngeal movement disorders are observed most often in individuals diagnosed with the neurological disorder, parkinsonism. Parkinsonism has the following features: bradykinesia, postural instability, rigidity, resting tremor, and freezing (motor blocks) (Fahn, 1986). For the diagnosis to be made, at least two of these five features should be present, and one of the two features should be either tremor or rigidity. Parkinsonism as a syndrome can be classified as idiopathic Parkinson’s disease (PD) (i.e., symptoms of unknown cause); secondary (or symptomatic) PD, caused by a known and identifiable cause; or parkinsonism-plus syndromes, in which symptoms of parkinsonism are caused by a known gene defect or have a distinctive pathology. The specific diagnosis depends on findings in the clinical history, the neurological examination, and laboratory tests. No single feature is completely reliable for differentiating among the different causes of parkinsonism.

Idiopathic PD is the most common type of parkinsonism encountered by the neurologist. Pathologically, idiopathic PD affects many structures in the central nervous system (CNS), with preferential involvement of dopaminergic neurons in the substantia nigra pars compacta (SNpc). Lewy bodies, eosinophilic intracytoplasmic inclusions, can be found in these neurons (Galvin, Lee, and Trojanowski, 2001). Alpha-synuclein is the primary component of Lewy body fibrils (Galvin, Lee, and Trojanowski, 2001). However, only about 75% of patients with the clinical diagnosis of idiopathic PD are found at autopsy to have the pathological CNS changes characteristic of PD (Hughes et al., 1992).

Many patients and their families consider the reduced ability to communicate one of the most difficult aspects of PD. Hypokinetic dysarthria, characterized by a soft voice, monotone, a breathy, hoarse voice quality, and imprecise articulation (Darley, Aronson, and Brown, 1975; Logemann et al., 1978), and reduced facial expression (masked facies) contribute to limitations in communication in the vast majority of individuals with idiopathic PD (Pitcairn et al., 1990). During the course of the disease, approximately 45%–89% of patients will report speech problems (Logemann and Fisher, 1981; Sapir et al., 2002). Repetitive speech phenomena (Benke et al., 2000), voice tremor, and hyperkinetic dysarthria may also be encountered in individuals with idiopathic PD. When hyperkinetic dysarthria is reported in idiopathic PD, it is most frequently seen together with other motor complications (e.g., dyskinesia) of prolonged levodopa therapy (Crichtley, 1981).

Logemann et al. (1978) suggested that the clusters of speech symptoms they observed in 200 individuals with PD represented a progression in dysfunction, beginning with disordered phonation in recently diagnosed patients and extending to include disordered articulation and other aspects of speech in more advanced cases. Recent findings by Sapir et al. (2002) are consistent with this
suggestion. Sapir et al. (2002) observed voice disorders in individuals with recent onset of PD and low Unified Parkinson Disease Rating Scale (UPDRS) scores; in individuals with longer duration of disease and higher UPDRS scores, they observed a significantly higher incidence of abnormal articulation and fluency, in addition to the disordered voice. Hypokinetic dysarthria of parkinsonism is considered to be a part of basal ganglia damage (Darley, Aronson, and Brown, 1975). However, there are no studies on pathological changes in the hypokinetic dysarthria of idiopathic PD. A significant correlation between neuronal loss and gliosis in SNpc and substantia nigra pars reticulata (SNpr) and severity of hypokinetic dysarthria was found in patients with Parkinson-plus syndromes (Kluin et al., 2001). Speech and voice characteristics may differ between idiopathic PD and Parkinson-plus syndromes (e.g., Shy-Drager syndrome, progressive supranuclear palsy, multisystem atrophy). In addition to the classic hypokinetic symptoms, these patients may have more slurring, a strained, strangled voice, palilalia, and hypernasality (Countryman, Ramig, and Pawlas, 1994) and their symptoms may progress more rapidly.

Certain aspects of hypokinetic dysarthria in idiopathic PD have been studied extensively. Hypophonia (reduced loudness, monotone, a breathy, hoarse quality) may be observed in as many as 89% of individuals with idiopathic PD (Logemann et al., 1978). Fox and Ramig (1997) reported that sound pressure levels in individuals with idiopathic PD were significantly lower (2–4 dB [30 cm]) across a variety of speech tasks than in an age- and sex-matched control group. Lack of vocal fold closure, including bowing of the vocal cords and anterior and posterior chinks (Hanson, Gerratt, and Ward, 1984; Smith et al., 1995), has been implicated as a cause of this hypophonia. Perez et al. (1996) used videostroboscopic observations to study vocal fold vibration in individuals with idiopathic PD. They reported abnormal phase closure and symmetry and tremor (both at rest and during phonation) in nearly 50% of patients. Whereas reduced loudness and disordered voice quality in idiopathic PD have been associated with glottal incompetence (lack of vocal fold closure—e.g., bowing; Hanson, Gerratt, and Ward, 1984; Smith et al., 1995; Perez et al., 1996), the specific origin of this glottal incompetence has not been clearly defined. R rigidity or fatigue secondary to rigidity, paralysis, reduced thyroarytenoid longitudinal tension secondary to cricothyroid rigidity (Aronson, 1990), and misperception of voice loudness (Ho, Bradshaw, and Iansek 2000; Sapir et al., 2002) are among the explanations. It has been suggested that glottal incompetence (e.g., vocal fold bowing) might be due to loss of muscle or connective tissue volume, either throughout the entire vocal fold or localized near the free margin of the vocal fold. Recent physiological studies of laryngeal function in idiopathic PD have shown a reduced amplitude of electromyographic activity in the thyroarytenoid muscle accompanying glottal incompetence when compared with both aged-matched and younger controls (Baker et al., 1998). These findings and the observation of reduced and variable single motor unit activity in the thyroarytenoid muscle of individuals with idiopathic PD (Luschei et al., 1999) are consistent with a number of hypotheses, the most plausible of which is reduced central drive to laryngeal motor neuron pools.

Although the origin of the hypophonia in PD is currently undefined, Ramig and colleagues (e.g., Fox et al., 2002) have hypothesized that there are at least three features underlying the voice disorder in individuals with PD: (1) an overall neural amplitude scaledown (Penny and Young, 1983) to the laryngeal mechanism (reduced amplitude of neural drive to the muscles of the larynx); (2) problems in sensory perception of effort (Berardelli et al., 1986), which prevents the individual with idiopathic PD from accurately monitoring his or her vocal output; which results in (3) the individual’s difficulty in independently generating (through internal cueing or scaling) adequate vocal effort (Hallet and Khoshbin, 1980) to produce normal loudness. Reduced neural drive, problems in sensory perception of effort, and problems scaling adequate vocal output effort may be significant factors underlying the voice problems in individuals with PD.

—Lorraine Olson Ramig, Mitchell F. Brin, Miodrag Velickovic, and Cynthia Fox

References


Fox, C., and Ramig, L. (1997). Vocal sound pressure level and self-perception of speech and voice in men and women with...


**Further Readings**


**Infectious Diseases and Inflammatory Conditions of the Larynx**

Infectious and inflammatory conditions of the larynx can affect the voice, swallowing, and breathing to varying extents. Changes can be acute or chronic and can occur in isolation or as part of systemic processes. The conditions described in this article are grouped by etiology.

**Infectious Diseases**

*Viral Laryngotraechitis.* Viral laryngotraechitis is the most common infectious laryngeal disease. It is typically associated with upper respiratory infection, for example, by rhinoviruses and adenoviruses. Dysphonia is usually self-limiting but may create major problems for a professional voice user. The larger diameter upper airway in
adults makes airway obstruction much less likely than in children. In a typical clinical scenario, a performer with mild upper respiratory symptoms has to carry on performing but complains of reduced vocal pitch and increased effort on singing high notes. Mild vocal fold edema and erythema may occur but can be normal for this patient group. Thickened, erythematous tracheal mucosa visible between the vocal folds supports the diagnosis. Hydration and rest may be sufficient treatment. However, if the performer decides to proceed with the show, high-dose steroids can reduce inflammation, and antibiotics may prevent opportunist bacterial infection. Cough suppressants, expectorants, and steam inhalations may also be useful. Careful vocal warmup should be undertaken before performing, and “rescue” must be balanced against the risk of vocal injury.

Other Viral Infections. Herpes simplex and herpes zoster infection have been reported in association with vocal fold paralysis (Flowers and Kernodle, 1990; Nishizaki et al., 1997). Laryngeal vesicles, ulceration, or plaques may lead to suspicion of the diagnosis, and antiviral therapy should be instituted early. New laryngeal muscle weakness may also occur in post-polio syndrome (Robinson, Hilllel, and Waugh, 1998). Viral infection has also been implicated in the pathogenesis of certain laryngeal tumors. The most established association is between human papillomavirus (HPV) and laryngeal papillomatosis (Levi et al., 1989). HPV, Epstein-Barr virus, and even herpes simplex virus have been implicated in the development of laryngeal malignancy (Ferlito et al., 1997; Garcia-Milian et al., 1998; Pou et al., 2000).

Bacterial Laryngitis. Bacterial laryngitis is most commonly due to Hemophilus influenzae, Staphylococcus aureus, Streptococcus pneumoniae, and beta-hemolytic streptococcus. Pain and fever may be severe, with airway and swallowing difficulties generally overshadowing voice loss. Typically the supraglottis is involved, with the aryepiglottic folds appearing boggy and edematous, often more so than the epiglottis. Unlike in children, laryngoscopy is usually safe in adults and is the best means of diagnosis. Possible underlying causes such as a laryngeal foreign body should be considered. Treatment includes intravenous antibiotics, hydration, humidification, and corticosteroids. Close observation is essential in case airway support is needed. Rarely, infected mucus retention cysts and epiglottic abscesses occur (Stack and Ridley, 1995). Tracheostomy and drainage may be required.

Mycobacterial Infections. Laryngeal tuberculosis is rare in industrialized countries but must be considered in the differential diagnosis of laryngeal disease, especially in patients with AIDS or other immune deficiencies (Singh et al., 1996). Tuberculosis can infect the larynx primarily, by direct spread from the lungs, or by hematogenous or lymphatic dissemination (Ramandan, Tarayi, and Baroudy, 1993). Most patients have hoarseness and odynophagia, typically out of proportion to the size of the lesion. However, these symptoms are not universally present. The vocal folds are most commonly affected, although all areas of the larynx can be involved. Laryngeal tuberculosis is often difficult to distinguish from carcinoma on laryngoscopy. Chest radiography and the purified protein derivative (PPD) test help establish the diagnosis, although biopsy and histological confirmation may be required. Patients are treated with antituberculous chemotherapy. The laryngeal symptoms usually respond within 2 weeks.

Leprosy is rare in developing countries. Laryngeal infection by Mycobacterium leprae can cause nodules, ulceration, and fibrosis. Lesions are often painless but may progress over the years to laryngeal stenosis. Treatment is with antileprosy chemotherapy (Soni, 1992).

Other Bacterial Infections. Laryngeal actinomycosis can occur in immunocompromised patients and following laryngeal radiotherapy (Nelson and Tybor, 1992). Biopsy may be required to distinguish it from radiocarcinoma or tumor. Treatment requires prolonged antibiotic therapy.

Scleroma is a chronic granulomatous disease due to Klebsiella scleromatis. Primary involvement is in the nose, but the larynx can also be affected. Subglottic stenosis is the main concern (Amoils and Shindo, 1996).

Fungal Laryngitis. Fungal laryngitis is rare and typically occurs in immunocompromised individuals. Fungi include yeasts and molds. Yeast infections are more frequent in the larynx, with Candida albicans most commonly identified (Vrabec, 1993). Predisposing factors in nonimmunocompromised patients include antibiotic and inhaled steroid use, and foreign bodies such as silicone voice prostheses.

The degree of hoarseness in laryngeal candidiasis may not reflect the extent of infection. Pain and associated swallowing difficulty may be present. Typically, thick white exudates are seen, and oropharyngeal involvement can coexist. Biopsy may show epithelial hyperplasia with a pseudocarcinomatous appearance. Potential complications include scarring, airway obstruction, and systemic dissemination.

In mild localized disease, topical nystatin or clotrimazole are usually effective. Discontinuing antibiotics or inhaled steroids should be considered. More severe cases may require oral antifungal azoles such as ketoconazole, fluconazole, or itraconazole. Intravenous amphotericin is efficacious but has potentially severe side effects. It is usually used for invasive or systemic disease.

Less common fungal diseases include blastomycosis, histoplasmosis, and coccidiomycosis. Infection may be confused with laryngeal carcinoma, and special histological stains are usually required for diagnosis. Long-term treatment with amphotericin B may be necessary.

Syphilis. Syphilis is caused by the spirochete Treponema pallidum. Laryngeal involvement is rare but may occur in later stages of the disease. Secondary syphilis
Chronic Laryngitis. Chronic laryngeal inflammation can result from smoking, gastroesophageal reflux (GER), voice abuse, or allergy. Patients often complain of hoarseness, sore throat, a globus sensation, and throat clearing. The vocal folds are usually thickened, dull, and erythematous. Posterior laryngeal involvement usually suggests GER. Besides direct chemical irritation, GER can promote laryngeal muscle misuse, which contributes to wear-and-tear injury (Gill and Morrison, 1998).

Inflammatory Processes

Chronic Laryngitis. Chronic laryngeal inflammation can result from smoking, gastroesophageal reflux (GER), voice abuse, or allergy. Patients often complain of hoarseness, sore throat, a globus sensation, and throat clearing. The vocal folds are usually thickened, dull, and erythematous. Posterior laryngeal involvement usually suggests GER. Besides direct chemical irritation, GER can promote laryngeal muscle misuse, which contributes to wear-and-tear injury (Gill and Morrison, 1998).

Although seasonal allergies may cause vocal fold edema and hoarseness (Jackson-Menaldi, Dzul, and Holland, 1999), it is surprising that allergy-induced chronic laryngitis is not more common. Even patients with significant nasal allergies or asthma have a low incidence of voice problems. The severity of other allergic accompaniments helps the clinician identify patients with dysphonia of allergic cause.

Treatment of chronic laryngitis includes voice rest and elimination of irritants. Dietary modifications and postural measures such as elevating the head of the bed can reduce GER. Proton pump inhibitors can be effective for persistent laryngeal symptoms (Hanson, Kamel, and Kahrilas, 1995).

Traumatic and Iatrogenic Causes. Inflammatory polyps, polypoid degeneration, and contact granuloma can arise from vocal trauma. Smoking contributes to polypoid degeneration, and intubation injury can cause contact granulomas. GER may promote inflammation in all these conditions. Granulomas can also form many years after Teflon injection for glottic insufficiency.

Rheumatoid Arthritis and Systemic Lupus Erythematosus. Laryngeal involvement occurs in almost a third of patients with rheumatoid arthritis (Lofgren and Montgomery, 1962). Patients present with a variety of symptoms. In the acute phase the larynx may be tender and inflamed. In the chronic phase the laryngeal mucosa may appear normal, but cricoarytenoid joint ankylosis may be present. Submucosal rheumatoid nodules or “bamboo nodes” can form in the membranous vocal folds. If the mucosal wave is severely damped, micro-laryngeal excision can improve the voice. Corticosteroids can be injected intracordally following excision. Other autoimmune diseases such as systemic lupus erythematosus can cause similar laryngeal pathology (Woo, Mendelsohn, and Humphrey, 1995).

Relapsing Polychondritis. Relapsing polychondritis is an autoimmune disease causing inflammation of cartilaginous structures. The pinna is most commonly affected, although laryngeal involvement occurs in around 50% of cases. Dapsone, corticosteroids, and immunosuppressive drugs have been used to control the disease. Repeated attacks of laryngeal chondritis can cause subglottic scarring, necessitating permanent tracheostomy (Spraggs, Tostevin, and Howard, 1997).

Cicatricial Pemphigoid. This chronic subepithelial bullous disease predominantly involves the mucous membranes. Acute laryngeal lesions are painful, and examination shows mucosal erosion and ulceration. Later, scarring and stenosis may occur, with supraglottic involvement (Hanson, Olsen, and Rogers, 1988). Treatment includes dapsone, systemic or intralesional steroids, and cyclophosphamide. Scarring may require laser excision and sometimes tracheostomy.

Amyloidosis. Amyloidosis is characterized by deposition of acellular proteinaceous material (amyloid) in tissues (Lewis et al., 1992). It can occur primarily or secondary to other diseases such as multiple myeloma or tuberculosis. Deposits may be localized or generalized. Laryngeal involvement is usually due to primary localized disease. Submucosal deposits may affect any part of the larynx but most commonly occur in the ventricular folds. Treatment is by conservative laser excision. Recurrences are frequent.

Sarcoidosis. Sarcoidosis is a multiorgan granulomatous disease of unknown etiology. About 6% of cases involve the larynx, producing dysphonia and airway obstruction. Pale, diffuse swelling of the epiglottis and aryepiglottic folds is characteristic (Benjamin, Dalton, and Croxson, 1995). Systemic or intralesional steroids, antilepromatous therapy, and laser debulking are all possible treatments.

Wegener’s Granulomatosis. Wegener’s granulomatosis is an idiopathic syndrome characterized by vasculitis and necrotizing granulomas of the respiratory tract and kidneys. The larynx is involved in 8% of cases (Waxman and Bose, 1986). Ulcerative lesions and subglottic stenosis may occur, causing hoarseness and dyspnea. Treatment includes corticosteroids and cyclophosphamide. Laser resection or open surgery is sometimes necessary for airway maintenance.

—David P. Lau and Murray D. Morrison

References


Further Readings


Instrumental Assessment of Children’s Voice

Disorders of voice may affect up to 5% of children, and instrumental procedures such as acoustics, aerodynamics, or electroglottography (EGG) may complement auditory-perceptual and imaging procedures by providing objective measures that help in determining the nature and severity of laryngeal pathology. The use of...
these procedures should take into account the developmental features of the larynx and special problems associated with a pediatric population.

An important starting point is the developmental anatomy and physiology of the larynx. This background is essential in understanding children’s vocal function as determined by instrumental assessments. The larynx of the infant and young child differs considerably in its anatomy and physiology from the adult larynx (see Anatomy of the Human Larynx). The vocal folds in an infant are about 3–5 mm long, and the composition of the folds is uniform. That is, the infant’s vocal folds are not only very short compared with those of the adult, but they lack the lamination seen in the adult folds. The lamination has been central to modern theories of phonation, and its absence in infants and marginal development in young children presents interesting challenges to theories of phonation applied to a pediatric population. An early stage of development of the lamina propria begins between 1 and 4 years, with the appearance of the vocal ligament (intermediate and deep layers of the lamina propria). During this same interval, the length of the vocal fold increases (reaching about 7.5 mm by age 5) and the entire laryngeal framework increases in size. The differentiation of the superficial layer of the lamina propria apparently is not complete until at least the age of 12 years.

Studies on the time of first appearance of sexual dimorphism in laryngeal size are conflicting, ranging from 3 years to no sex differences in laryngeal size observable during early childhood. Sexual dimorphism of vocal fold length has been reported to appear at about age 6–7 years. These reported anatomical differences do not appear to contribute to significant differences in vocal fundamental frequency ($f_0$) between males and females until puberty, at which time laryngeal growth is remarkable, especially in boys. For example, in boys, the anteroposterior dimension of the thyroid cartilage increases threefold, along with increases in vocal fold length.

**Acoustic Studies of Children’s Voice.** Mean $f_0$ has been one of the most thoroughly studied aspects of the pediatric voice. For infants’ nondistress utterances, such as cooing and babbling, mean $f_0$ falls in the range of 300–600 Hz and appears to be stable until about 9 months, when it begins to decline until adulthood (Kent and Read, 2002). A relatively sharp decline occurs between the ages of 12 months and 3 years, so that by the age of 3 years, the mean $f_0$ in both males and females is about 250 Hz. Mean $f_0$ is stable or gradually falling between 6 and 11 years, and the value of 250 Hz may be taken as a reasonable estimate of $f_0$ in both boys and girls. Some studies report no significant change in $f_0$ during this developmental period, but Glaze et al. (1988) reported that $f_0$ decreased with increasing age, height, and weight for boys and girls ages 5–11 years, and Ferrand and Bloom (1996) observed a decrease in the mean, maximum, and range of $f_0$ in boys, but not in girls, at about 7–8 years of age.

Sex differences in $f_0$ emerge especially strongly during adolescence. The overall $f_0$ decline from infancy to adulthood is about one octave for girls and two octaves for boys. There is some question as to when the sex difference emerges. Lee et al. (1999) observed that $f_0$ differences between male and female children were statistically significant beginning at about age 12 years, but Glaze et al. (1988) observed differences between boys and girls for the age period 5–11 years. Further, Hacki and Heitmuller (1999) reported a lowering of both the habitual pitch and the entire speaking pitch range between the ages of 7 and 8 years for girls and between the ages of 8 and 9 years for boys. Sex differences emerge strongly with the onset of mutation. Hacki and Heitmuller (1999) concluded that the beginning of the mutation occurs at age 10–11 years. Mean $f_0$ change is pronounced in males between the ages of about 12 and 15 years. For example, Lee et al. (1999) reported a 78% decrease in $f_0$ for males between these ages. No significant change was observed after the age of 15 years, which indicates that the voice change is effectively complete by that age (Hollien, Green, and Massey, 1994; Kent and Vorperian, 1995).

Other acoustic aspects of children’s voices have not been extensively studied. In apparently the only large-scale study of its kind, Campisi et al. (2002) provided normative data for children for the parameters of the Multi-Dimensional Voice Program (MDVP). On the majority of parameters (excluding, of course, $f_0$), the mean values for children were fairly consistent with those for adults, which simplifies the clinical application of MDVP. However, this conclusion does not apply to the pubescent period, during which variability in amplitude and fundamental frequency increases in both girls and boys, but markedly so in the latter (Boltezar, Burger, and Zargi, 1997). It should also be noted voice training can affect the degree of aperiodicity in children’s voices (Dejonckere et al., 1996) (see Acoustic Assessment of Voice).

**Aerodynamic Studies of Children’s Voice.** There are only limited data describing developmental patterns in voice aerodynamics. Table 1 shows normative data for flow, pressure, and laryngeal airway resistance from three sources (Netsell et al., 1994; Keilman and Bader, 1995; Zajac, 1995, 1998). All of the data were collected during the production of /pi/ syllable trains, following the procedure first described by Smitheran and Hixon (1981). Flow appears to increase with age, ranging from 75–79 mL/s in children aged 3–5 years to 127–188 mL/s in adults. Pressure decreases slightly with age, ranging from 8.4 cm H$_2$O in children aged 3–5 years to 5.3–6.0 cm H$_2$O in adults. Laryngeal airway pressure decreases with age, ranging from 111–119 cm H$_2$O/L/s in children aged 3–6 years to 34–43 cm H$_2$O/L/s in adults. This decrease in laryngeal airway pressure occurs as a function of the rate of flow increase exceeding the rate of pressure decrease across the age range.

Netsell et al. (1994) explained the developmental changes in flow, pressure, and laryngeal airway pressure...
Table 1. Aerodynamic normative data from three sources: N (Netsell et al., 1994), K (Keilman & Bader, 1995), and Z (Zajac, 1995, 1998). All data were collected using the methodology described by Smitheran and Hixon (1981). Values shown are means, with standard deviations in parentheses.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Age (yr)</th>
<th>Sex</th>
<th>N</th>
<th>Flow (mL/s)</th>
<th>Pressure (cm H₂O)</th>
<th>LAR (cm H₂O/L/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>3–5</td>
<td>F</td>
<td>10</td>
<td>79 (16)</td>
<td>8.4 (1.3)</td>
<td>111 (26)</td>
</tr>
<tr>
<td>N</td>
<td>3–5</td>
<td>M</td>
<td>10</td>
<td>75 (20)</td>
<td>8.4 (1.4)</td>
<td>119 (20)</td>
</tr>
<tr>
<td>K</td>
<td>4–7</td>
<td>F&amp;M</td>
<td>10</td>
<td>7.46 (2.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>6–9</td>
<td>F</td>
<td>10</td>
<td>86 (19)</td>
<td>7.4 (1.5)</td>
<td>89 (25)</td>
</tr>
<tr>
<td>N</td>
<td>6–9</td>
<td>M</td>
<td>9</td>
<td>101 (42)</td>
<td>8.3 (2.0)</td>
<td>97 (39)</td>
</tr>
<tr>
<td>Z</td>
<td>7–11</td>
<td>F&amp;M</td>
<td>10</td>
<td>123 (30)</td>
<td>11.4 (2.3)</td>
<td>95.3 (24.4)</td>
</tr>
<tr>
<td>K</td>
<td>8–12</td>
<td>F&amp;M</td>
<td>10</td>
<td>6.81 (2.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>9–12</td>
<td>F</td>
<td>10</td>
<td>121 (21)</td>
<td>7.1 (1.2)</td>
<td>59 (7)</td>
</tr>
<tr>
<td>N</td>
<td>9–12</td>
<td>M</td>
<td>10</td>
<td>115 (42)</td>
<td>7.9 (1.3)</td>
<td>77 (23)</td>
</tr>
<tr>
<td>K</td>
<td>13–15</td>
<td>F&amp;M</td>
<td>100</td>
<td>50–150</td>
<td>5.97 (2.07)</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>4–15</td>
<td>F&amp;M</td>
<td>100</td>
<td>50–150</td>
<td>5.97 (2.07)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Adult</td>
<td>F</td>
<td>10</td>
<td>127 (29)</td>
<td>5.3 (1.2)</td>
<td>43 (10)</td>
</tr>
<tr>
<td>N</td>
<td>Adult</td>
<td>M</td>
<td>10</td>
<td>188 (51)</td>
<td>6.0 (1.4)</td>
<td>34 (9)</td>
</tr>
</tbody>
</table>

F = female, M = male, N = number of participants, LAR = laryngeal airway resistance.

Instrumental Assessment of Children’s Voice

as secondary to an increasing airway size and decreasing dependence on expiratory muscle forces alone for speech breathing with age. No consistent differences in aerodynamic parameters were observed between female and male children. High standard deviations reflect considerable variation between children of similar ages (see AERODYNAMIC ASSESSMENT OF VOICE).

Electroglottographic Studies of Children’s Voice. Although EGG data on children’s voice are not abundant, one study provides normative data on a sample of 164 children, 79 girls and 85 boys, ages 3–16 years (Cheyne, Nuss, and Hillman, 1999). Cheyne et al. reported no significant effect of age on the EGG measures of jitter, open quotient, closing quotient, and opening quotient. The means and standard deviations (in parentheses) for these measures were as follows: jitter—0.76% (0.61), open quotient—54.8% (3.3), closing quotient—14.1% (3.8), and opening quotient—31.1% (4.1). These values are reasonably similar to values reported for adults, although caution should be observed because of differences in procedures across studies (Takahashi and Koike, 1975) (see ELECTROGLOTTOGRAPHIC ASSESSMENT OF VOICE).

One of the most striking features of the instrumental studies of children’s voice is that, except for \( f_0 \) and the aerodynamic measures, the values obtained from instrumental procedures change relatively little from childhood to adulthood. This stability is remarkable in view of the major changes that are observed in laryngeal anatomy and physiology. Apparently, children are able to maintain normal voice quality in the face of considerable alteration in the apparatus of voice production. With the mutation, however, stability is challenged, and the suitability of published normative data is open to question. The maintenance of rather stable values across a substantial period of childhood (from about 5 to 12 years) for many acoustic and EGG parameters holds a distinct advantage for clinical application. It is also clear that instrumental procedures can be used successfully with children as young as 3 years of age. Therefore, these procedures may play a valuable role in the objective assessment of voice in children.

See also VOICE DISORDERS IN CHILDREN.

Ray D. Kent and Nathan V. Welham

References


Netsell, R., Lotz, W. K., Peters, J. E., and Schulte, L. (1994). Tremor, and paradoxical breathing dystonia. All of these conditions are idiopathic and all have distinctive symptoms, which form the basis for diagnosis. In adductor spasmodic dysphonia (ADSD), voice breaks during vowels are associated with involuntary spasmodic muscle bursts in the thyroarytenoid and other adductor laryngeal muscles, although bursts can also occur in the cricothyroid muscle in some persons (Nash and Ludlow, 1996). When voice breaks are absent, however, muscle activation is normal in both adductor and abductor laryngeal muscles (Van Pelt, Ludlow, and Smith, 1994). In the abductor type of spasmodic dysphonia (ABSD), breathy breaks are due to prolonged vocal fold opening during voiceless consonants (p, t, k, f, s, h), which are perceived as breathy breaks, in ASBD. Other idiopathic voice disorders, such as muscular tension dysphonia, do not involve intermittent spasmodic changes in the voice. Rather, consistent abnormal hypertense laryngeal postures are maintained during voice production. Such persons may respond to manual laryngeal manipulation (Roy, Ford, and Bles, 1996).

Muscular tension dysphonia may be confused with spasmodic dysphonia when ADSD patients develop increased muscle tension in an effort to overcome vocal instability, resulting in symptoms of both disorders. Some patients with voice tremor may also develop muscular tension dysphonia in an effort to overcome vocal instability. Paradoxical breathing dystonia is rare, with adduc- tory movements of the vocal folds during inspiration that remit during sleep (Marion et al., 1992). It differs from vocal fold dysfunction, which is usually intermittent and often coincides with irritants affecting the upper airway (Christopher et al., 1983; Morrison, Rammage, and Emami, 1999).

Botulinum toxin type A (BTX-A) is effective in treating a myriad of hyperkinetic disorders by partially denervating the muscle. The toxin is injected into muscle, diffuses, and is endocytosed into nerve endings. The toxin cleaves SNAP 25, a vesicle-docking protein essential for acetylcholine release into the neuromuscular junction (Aoki, 2001). When acetylcholine release is blocked, the muscle fibers become temporarily denervated. The effect is reversible: within a few weeks new nerve endings sprout, which may provide synaptic transmission and some reduction in muscle weakness. These nerve endings are later replaced by restitution of the original end-plates (de Paiva et al., 1999). In ADSD, BTX-A injection, either small bilateral injections or a unilateral injection produces a partial chemodenervation of the thyroarytenoid muscle for up to 4 months. Reductions in spasmodic muscle bursts relate to voice improvement (Bielamowicz and Ludlow, 2000). This therapy is effective in at least 90% of ADSD patients, as has been demonstrated in a small randomized controlled trial (Truong et al., 1991) and in multiple case series (Ludlow et al., 1988; Blitzer and Brin, 1991; Blitzer, Brin, and Stewart, 1998). BTX-A is less effective in ABSD. When only ABSD patients with cricothyroid muscle spasms are injected in that muscle, significant improvements occur in 60% of cases (Ludlow et al., 1991). Similarly, two-thirds of ABSD patients obtain some degree of benefit from posterior cricoarytenoid

**Laryngeal Movement Disorders: Treatment with Botulinum Toxin**

The laryngeal dystonias include spasmodic dysphonia, tremor, and paradoxical breathing dystonia. All of these conditions are idiopathic and all have distinctive symptoms, which form the basis for diagnosis. In adductor spasmodic dysphonia (ADSD), voice breaks during vowels are associated with involuntary spasmodic muscle bursts in the thyroarytenoid and other adductor laryngeal muscles, although bursts can also occur in the cricothyroid muscle in some persons (Nash and Ludlow, 1996). When voice breaks are absent, however, muscle activation is normal in both adductor and abductor laryngeal muscles (Van Pelt, Ludlow, and Smith, 1994). In the abductor type of spasmodic dysphonia (ABSD), breathy breaks are due to prolonged vocal fold opening during voiceless consonants. The posterior cricoarytenoid muscle is often involved in ABSD, although not in all patients (Cyrus et al., 2001). In the 1980s, “spastic” dysphonia was renamed “spasmodic” dysphonia to denote the intermittent aspect of the voice breaks and was classified as a task-specific focal laryngeal dystonia (Blitzer and Brin, 1991). Abnormalities in laryngeal adductor responses to sensory stimulation are found in both ADSD and ABSD (Deleyiannis et al., 1999), indicating a reduction in the normal central suppression of laryngeal sensorimotor responses in these disorders. ADSD affects 85% of patients with spasmodic dysphonia; the other 15% have ABSD.

Vocal tremor is present in at least one-third of patients with ADSD or ABSD and can also occur in isolation. A 5-Hz tremor can be heard on prolonged vowels, owing to intensity and frequency modulation. Tremor can affect either or both the adductor or abductor muscles, producing voice breaks in vowels or breathy intervals in the abductor type. Voice tremor occurs more often in women, sometimes with an associated head tremor. A variety of muscles may be involved in voice tremor (Koda and Ludlow, 1992).

Intermittent voice breaks are specific to the spasmodic dysphonias, either prolonged glottal stops and intermittent intervals of a strained or strangled voice quality during vowels in ADSD or prolonged voiceless consonants (p, t, k, f, s, h), which are perceived as breathy breaks, in ABSD. Other idiopathic voice disorders, such as muscular tension dysphonia, do not involve intermittent spasmodic changes in the voice. Rather, consistent abnormal hypertense laryngeal postures are maintained during voice production. Such persons may respond to manual laryngeal manipulation (Roy, Ford, and Bles, 1996). Muscular tension dysphonia may be confused with spasmodic dysphonia when ADSD patients develop increased muscle tension in an effort to overcome vocal instability, resulting in symptoms of both disorders. Some patients with voice tremor may also develop muscular tension dysphonia in an effort to overcome vocal instability.
injections (Blitzer et al., 1992). When speech symptoms were measured in blinded fashion before and after treatment, BTX-A was less effective in ABSD (Bielamowicz et al., 2001) than in ADSD (Ludlow et al., 1988).

Patients with adductor tremor confined to the vocal folds often receive some benefit from thyroarytenoid muscle BTX-A injections. When objective measures were used (Warrick et al., 2000), BTX-A injection was beneficial in 50% of patients with voice tremors. Either unilateral or bilateral thyroarytenoid injections can be used, although larger doses are sometimes more effective. BTX-A is much less effective, however, for treating tremor than it is in ADSD (Warrick et al., 2000), and it is rarely helpful in patients with abductor tremor.

BTX-A administered as either unilateral or bilateral injections into the thyroarytenoid muscle has been used successfully to treat paradoxical breathing dystonia (Marion et al., 1992; Grillone et al., 1994).

Changes in laryngeal function following BTX-A injection in persons with ADSD are similar, whether the injection was unilateral or bilateral. A few persons report a sense of reduction in laryngeal tension within 8 hours following injection, although voice loudness is not yet reduced. Voice loudness and breaks gradually diminish as BTX-A diffuses through the muscle, causing progressive denervation. Most people report that benefits become apparent the second day, while the side effects of progressive breathiness and swallowing difficulties increase over the 3–5 days after injection. Difficulty swallowing liquids may occur and occasionally results in aspiration. Patients are advised to ingest liquids slowly and in small volumes, by sipping through a straw. The difficulties with swallowing gradually subside between the first and second weeks after an injection (Ludlow, Rhew, and Nash, 1994), possibly as patients learn to compensate. The breathiness resolves somewhat later, reaching normal loudness levels as late as 3–4 weeks after injection, during the period when axonal sprouting may occur (de Paiva et al., 1999). Because improvements in voice volume seem independent of recovery of swallowing, different mechanisms may underlie recovery of these functions.

The benefit is greatest between 1 and 3 months after injection, when the patient’s voice is close to normal volume, voice breaks are significantly reduced, and speech is more fluent, with reduced hoarseness (Ludlow et al., 1988). This benefit period differs among the disorders, lasting from 3 to 5 months in ADSD but from 1 to 3 months in other disorders such as ABSD and tremor.

The return of symptoms in ADSD is gradual, usually occurring over a period of about 2 months during endplate reinnervation. To maintain symptom control, most patients return for injection about 3 months before the full return of symptoms. Some individuals, however, maintain benefit for more than a year following injection, returning 2 or more years later for reinjection.

The mechanisms responsible for benefit from BTX-A in laryngeal dystonia likely differ with the different pathophysiology: although BTX-A is beneficial in many hyperkinetic disorders, it is more effective in some than in others. In all cases BTX-A causes partial denervation, which reduces the force that can be exerted by a muscle following injection. In ADSD, then, the forcefulness of vocal fold hyperadduction is reduced and patients are less able to produce voice breaks even if muscle spasms continue to occur. Central control changes also appear to occur, however, in persons with ADSD following BTX-A injection. When thyroarytenoid muscle injections were unilateral, spasmodic bursts were significantly reduced on both sides of the larynx, and there were also reductions in overall levels of muscle tone (measured in microvolts) and maximum activity levels (Bielamowicz and Ludlow, 2000). Such reductions in muscle activity and spasms may be the result of reductions in muscle spindle and mechanoreceptor feedback, resulting in lower motor neuron pool activity for all the laryngeal muscles. The physiological effects of BTX-A may be greater on the fusimotor system than on muscle fiber innervation (On et al., 1999). Although only one portion of the human thyroarytenoid muscle may contain muscle spindles (Sanders et al., 1998), mucosal mechanoreceptor feedback will also change with reductions in adductory force between the vocal folds following BTX-A injection. Perhaps changes in sensory feedback account for the longer period of benefit in ADSD than in other laryngeal movement disorders, although the duration of side effects is similar in all disorders. Future approaches to altering sensory feedback may also have a role in the treatment of laryngeal dystonia, in addition to efferent denervation by BTX-A.

—Christy L. Ludlow

References


**Further Readings**


Laryngeal Reinnervation Procedures

The human larynx is a neuromuscularly complex organ responsible for three primary and often opposing functions: respiration, swallowing, and speech. The most primitive responsibilities include functioning as a conduit to bring air to the lungs and protecting the respiratory tract during swallowing. These duties are physiologically opposite; the larynx must form a wide caliber during respiration but also be capable of forming a tight sphincter during swallowing. Speech functions are fine-tuned permutations of laryngeal opening and closure against pulmonary airflow. Innervation of this organ is complex, and its design is still being elucidated. Efferent fibers to the larynx from the brainstem motor nuclei travel by way of the vagus nerve to the superior laryngeal nerve (SLN) and the recurrent laryngeal nerves (RLNs). Afferent fibers emanate from intramuscosal and intramuscular receptors and travel along pathways that include the SLN (supraglottic larynx) and the RLN (subglottic larynx). Autonomic fibers also innervate the larynx, but these are poorly understood (see also Anatomy of the Human Larynx).

Reinnervation of the larynx was first reported 1909 by Horsley, who described reanastomosis of a severed RLN. Reports of laryngeal reinnervation spotted the literature over the next several decades, but it was not until the last 30 years that reinnervation techniques were refined and became performed with relative frequency. Surely this is associated with advances in surgical optics as well as microsurgical instrumentation and technique. Indications for laryngeal reinnervation include functional reanimation of the paralyzed larynx, prevention of denervation atrophy, restoration of laryngeal sensation, and modification of pathological innervation (Crumley, 1991; Aviv et al., 1997; Berke et al., 1999).

Anatomy and Technique of Reinnervation. The technique of reinnervation is similar for both sensory and motor systems. The nerve in question is identified through a transcervical approach. Usually a horizontal skin incision is placed into a neck skin crease at about the level of the cricoid cartilage. Subplatysmal flaps are elevated and retracted. The larynx is further exposed after splitting the strap muscles in the midline. Sensation to the supraglottic mucosal is supplied via the internal (sensory) branch of the SLN. This structure is easily identified as it pierces the thyrohyoid membrane on either side (Fig. 1). The motor innervation of the larynx is somewhat more complicated. All of the intralaryngeal muscles are innervated by the RLN. This nerve approaches the larynx from below in the tracheoesophageal groove. The nerve enters the larynx from deep to the cricothyroid joint and immediately splits into an anterior and a posterior division (Fig. 2). The anterior division supplies all of the laryngeal adductors except the cricothyroid muscle; the posterior division supplies the only abductor of the...
larynx, the posterior cricoarytenoid muscle. The anterior division further arborizes to innervate each of the intrinsic adductors in a well-defined order: the interarytenoid followed by the lateral cricoarytenoid, and lastly the thyroarytenoid muscles. The interarytenoid muscle is thought to receive bilateral innervation, while the other muscles all receive unilateral innervation. The external branch of the SLN innervates the only external adductor of the larynx—the cricothyroid muscle. Although the SLN and the main trunk of the RLN can be approached without opening the larynx, the adductor branches of the RLN can only be successfully approached after opening the thyroid cartilage. A large inferiorly based window is made in the thyroid lamina and centered over the inferior tubercle. Once the cartilage is opened the anterior branch can be seen coursing obliquely toward the terminus in the midportion of the thyroarytenoid muscle. The posterior division is approached by rotation of the larynx, similar to an external approach to the arytenoid. Identification and dissection of the fine distal nerve branches is usually carried out under loupes or microscopic magnification using precision instruments.

Once the damaged nerve is identified, it is severed sharply with a single cut of a sharp instrument. This is important to avoid crushing trauma to the nerve stump. The new motor or sensory nerve is brought into the field under zero tension and then anastomosed with several epineurial sutures of fine microsurgical material (9-0 or 10-0 nylon or silk). The anastomosis must be tension-free. The donor nerve is still connected to its proximal (motor) or distal (sensory) cell bodies. Selection of the appropriate donor nerve is discussed subsequently.

Over the next 3–9 months healing occurs, with neurontization of the motor end-plates or the sensory receptors. Although physiological reinnervation is the goal, relatively few reports have demonstrated true volitional movement. Typically, one can hope to prevent muscle atrophy and help restore muscle bulk. Sensory reinnervation is less clear.

Prior to the microsurgical age, most reinnervation procedures of the larynx were carried out with nerve-muscle pedicle implantation into the affected muscle. With modern techniques, end-to-end nerve-nerve anastomosis with epineural suture fixation is a superior and far more reliable technique of reinnervation.

Reinnervation for Laryngeal Paralysis. Paralysis of the larynx is described as unilateral or bilateral. The diagnosis is made on the basis of history and physical examination including laryngoscopic findings, and sometimes on the basis of laryngeal electromyography or radiographic imaging. The unilaterally paralyzed larynx is more common and is characterized by a lateralized vocal cord that prevents complete glottic closure during laryngeal tasks. The patient seeks care for dysphonia and aspiration or cough during swallowing. The etiology is commonly idiopathic, but the condition may be due to inflammatory neuropathy, iatrogenic trauma, or neoplastic invasion of the recurrent nerve. Although many patients are successfully treated with various static procedures, one can argue that, theoretically, the best results would restore the organ to its preexisting physiological state. Over the past 20 years there has been increased interest in physiological restoration, a concept championed by Crumley (Crumley, 1983, 1984, 1991; Crumley, Izdebski, and McMicken, 1988). A series of patients with unilateral vocal cord paralysis were treated with anastomosis of the distal RLN trunk to the ansa cervicalis. The ansa cervicalis is a good example of an acceptible motor donor nerve (Crumley, 1991; Berke et al., 1999). This nerve normally supplies motor neurons to the strap muscles (extrinsic accessory muscles of
the larynx), whose function is to elevate and lower the larynx during swallowing. Fortunately, when the ansa cervicalis is sacrificed, the patient does not have noticeable disability. The size match to the RLN is excellent when using either the whole nerve bundle for main trunk anastomosis or an easily identifiable fascicle for connection to the anterior branch.

Hemilaryngeal reinnervation with the ansa cervicalis has been shown to improve voicing in those patients undergoing the procedure (Crumley and Izdbski, 1986; Chhetri et al., 1999). Evaluation of these patients, however, does not demonstrate restoration of normal muscular physiology. The affected vocal cord has good bulk, but volitional movement typically is not restored. Proponents of other techniques have argued that the ansa cervicalis may not have enough axons to properly regenerate the RLN, or that synkinesis has occurred (Paniello, Lee, and Dahm, 1999). Synkinesis refers to mass firing of a motor nerve that can occur after reinnervation. In the facial nerve, for example, one may see mass movement of the face with volitional movement because all the branches are essentially acting as one. The RLN contains both abductor and adductor (as well as a small amount of sensory and autonomic) fibers. With reinnervation, one may hypothesize that mass firing of all fibers cancels the firings of individual fibers out and thus produces a static vocal cord. With this concept in mind, some have recommended combining reinnervation with another static procedure to augment results (combination with arytenoid adduction) or to avoid the potential for synkinesis by performing the anastomosis of the donor nerve to the anterior branch of the recurrent nerve (Green et al., 1992; Nasri et al., 1994; Chhetri et al., 1999).

Paniello has proposed that the ansa cervicalis is not the best donor nerve to the larynx (Paniello, Lee, and Dahm, 1999). He suggests that the hypoglossal nerve would be more appropriate because of increased axon bulk and little donor morbidity. Animal experiments with this technique have demonstrated volitional and reflexive movement of the reinnervated vocal cord.

Neurological bilateral vocal cord paralysis is often post-traumatic or iatrogenic. These patients are troubled by a fixed small airway and often find themselves tracheotomy dependent. Therapy is directed at restoring airway caliber while avoiding aspiration. Voice issues are usually considered secondary to the airway concerns. Although most practitioners currently treat with static techniques, physiological restoration would be preferred. In the 1970s, Tucker advocated reinnervation of the posterior cricoarytenoid with a nerve muscle pedicle of the sternohyoid muscle and ansa cervicalis (Tucker, 1978). Although his results were supportive of his hypothesis, many have had trouble repeating them. More recently, European groups have studied phrenic nerve to posterior branch of the RLN transfers (van Lith-Bijl et al., 1997, 1998). The phrenic nerve innervates the diaphragm and normally fires with inspiration; it has been shown that one of its nerve roots can be sacrificed without paralysis of the diaphragm. Others have suggested use of the SLN as a motor source for the posterior cricoarytenoid muscle (Maniglia et al., 1989). Most techniques of reinnervation for bilateral vocal cord paralysis still have not enjoyed the success of unilateral reinnervation and are only performed by a few practitioners. The majority of patients with bilateral vocal cord paralysis undergo a static procedure such as cordotomy, arytenoidectomy, or tracheotomy to improve their airway.

**Sensory Palsy.** Recent work has highlighted the importance of laryngeal sensation. After development of an air-pulse quantification system to measure sensation, it was shown that patients with stroke and dysphagia have a high incidence of laryngosensory deficit. Studies performed in the laboratory demonstrated that reanastomosis of the internal branch of the SLN restored protective laryngeal reflexes (Blumin, Berke, and Blackwell, 1999). Clinically, anastomosis of the greater auricular nerve to the internal branch of the SLN has been successfully used to restore sensation to the larynx (Aviv et al., 1997).

**Modification of Dystonia.** Spasmodic dysphonia is an idiopathic focal dystonia of the larynx. The majority of patients have the adductor variety, characterized by intermittent and paroxysmal spasms of the vocal cords during connected speech. The mainstay of treatment for this disorder is botulinum toxin (Botox) injections into the affected laryngeal adductor muscles. Unfortunately, the effect of Botox is temporary, and repeated injections are needed indefinitely. A laryngeal denervation and reinnervation procedure has been designed to provide a permanent alternative to Botox treatment (Berke et al., 1999). In this procedure, the distalmost branches of the laryngeal adductors are severed from their muscle insertion. These “bad” nerve stumps are then sutured outside the larynx to avoid spontaneous reinnervation. A fascicle of the ansa cervicalis is then suture-anastomosed to the distal thyroarytenoid branch for reinnervation. Reinnervation maintains tone of the thyroarytenoid, thus preventing atrophy and theoretically protecting that muscle by occupying the motor end-plates with neurons unaffected by the dystonia. This approach has had great success, with about 95% of patients achieving freedom from further therapy.

**Laryngeal Transplantation.** Transplantation of a physiologically functional larynx is the sought-after grail of reinnervation. For a fully functional larynx, eight nerves would be anastomosed—bilateral anterior and posterior branches of the RLN and bilateral external and internal branches of the SLN. To date, success has been reliably achieved in the canine model (Berke et al., 1993) and partially achieved in one human (Strome et al., 2001). Current research has been aimed at preventing transplant rejection. The technique of microneural, microvascular, and mucosal anastomosis has been well worked out in the animal model.

—Joel H. Blumin and Gerald S. Berke
References


Further Readings


Alterations of the vibratory, articulatory, or resonance system consequent to traumatic injury or the treatment of disease can significantly alter the functions of both voice and speech, and potentially deglutition and swallowing. In some instances, subsequent changes to the larynx and oral peripheral system may be relatively minor and without substantial consequence to the individual. In other instances these changes may result in dramatic alteration of one or more anatomical structures necessary for normal voice and speech production, in addition to other oral functions. Traumatic injury and surgical treatment for disease also may affect isolated structures of the peripheral speech mechanism, or may have more widespread influences on entire speech subsystems (e.g., articulatory, velopharyngeal) and the related structures necessary for competent and effective verbal communication.

Laryngeal Trauma

Trauma to the larynx is a relatively rare clinical entity. Estimates in the literature indicate that acute laryngeal trauma accounts for 1 in 15,000 to 1 in 42,000 emergency room visits. A large number of such traumatic injuries are the result of accidental blunt trauma to the neck; causes include motor vehicle collisions, falls, athletic injuries, and the like. Another portion of these injuries are the result of violence, such as shooting or stabbing, which may result in penetrating injuries not only of the larynx but also of other critical structures in the neck. Blunt laryngeal trauma is most commonly reported in persons less than 30 years old.

In cases of blunt trauma to the larynx, the primary presenting symptoms are hoarseness, pain, dyspnea (shortness of breath), dysphagia, and swelling of the tissues of the neck (cervical emphysema). Injuries may involve fractures of laryngeal cartilages, partial or complete dislocations, lacerations of soft tissues, or combined types of injury. Because laryngeal trauma, no matter how minor, holds real potential to affect breathing, medical intervention is first directed at determining airway patency and, when necessary, maintaining an adequate airway through emergency airway management (Schaefer, 1992). When airway compromise is observed, emergency tracheotomy is common. Laryngeal trauma is truly an emergency medical condition. When injuries are severe, additional surgical treatment may be warranted. Thus, whereas vocal disturbances are possible, the airway is of primary concern; changes to the voice are of secondary importance.

Injuries to the intrinsic structures of the oral cavity are also rare, although when they do occur, changes in speech, deglutition, and swallowing may exist. Although the clinical literature is meager in relation to injuries of the lip, alveolus, floor of the mouth, tongue, hard palate, and velum, such injuries or their medical treatment may have a significant impact on speech. Trauma to the mandible also can directly impact verbal communication. Unfortunately, the literature in this area is sparse, and information on speech outcomes following injuries of this type is frequently anecdotal. However, in addressing any type of traumatic injury to the peripheral structures of the speech mechanism, assessment methods typically employed with the dysarthrias may be most appropriate (e.g., Dworkin, 1991). In this regard, the point-place system may provide essential information on the extent and degree of impairment of speech subsystems (Netsell and Daniel, 1979).

Speech Considerations. Speech management initially focuses on identifying which subsystems are impaired, the severity of impairment, and the consequent reduction in speech intelligibility and communicative proficiency. A comprehensive evaluation that involves aerodynamic, acoustic, and auditory-perceptual components is essential. Information from each of these areas is valuable in identifying the problem, developing management strategies, and monitoring patient progress. Because of variability in the extent of traumatic injuries, the literature on the dysarthrias often provides a useful framework for establishing clinical goals and evaluating treatment effectiveness.

Peripheral Structural Changes Resulting from Surgical Ablation

In contrast to peripheral structural changes due to traumatic injury, structural changes due to surgical ablation for oral cancer also result in alterations in functions necessary for speech or swallowing. The treatment itself has clear potential to affect speech production, but changes in speech may be variable and ultimately depend on the structures treated.

Malignant Conditions Affecting Peripheral Structures of the Speech System. Head and neck squamous cell carcinoma may occur in the epithelial tissue of the mouth, nasal cavity, pharynx, or larynx. Tumors of the head and neck account for approximately 5% of all malignancies in men and 2% of all malignancies in women (Franceschi et al., 2000). The majority of head and neck cancers involve squamous epithelial tissue. These tumors have the potential to be invasive; therefore, the potential for spread of disease is substantial, and radical dissection of regional lymphatics is often required. Because of the possibility of disease spread, radiation therapy is commonly used to eliminate occult disease, and this treatment may have negative consequences on vocal and speech functions.

As tumors increase in size, they may invade adjacent tissue, which frequently requires more extensive...
treatment because of the threat of distant spread of disease. Some individuals are asymptomatic at the time of diagnosis; others report pain, difficulty swallowing (dysphagia), or difficulty breathing. The destructive nature of a malignant tumor may cause adjacent muscular structures of the tongue or floor of mouth to become fixed. In other situations, the tumor may encroach on the nerve supply, which is likely to result in loss of sensation (paresthesia), paralysis, or pain. In both circumstances, the potential for metastatic spread of disease is substantial.

Current treatment protocols for tumors involving structures of the peripheral speech mechanism include surgery, radiotherapy, or chemotherapy, alone or in combination. The choice of modality usually depends on tumor location, disease stage, cell type, and other factors. Each treatment modality is associated with some additional morbidity that can significantly affect structure, function, cosmesis, and quality of life. Surgery carries a clear potential for anatomical and physiological changes that may directly alter speech and swallowing, while at the same time creating significant cosmetic deformities. Similarly, radiotherapy is commonly associated with a range of side effects. Radiation delivered to the head and neck affects both normal and normal tissues. Salivary glands may be damaged, with a resulting decrease in salivary flow leading to a dry mouth (xerostomia). This decrease in saliva may then challenge normal oral hygiene and health, which may result in dental caries following radiation treatment. Osteoradionecrosis may result in the exposure of bone, tissue changes, or infection, which usually cause significant discomfort to the individual. Osteoradionecrosis may be decreased if the radiation exposure is limited to less than 60 cGy (Thorn et al., 2000). In some cases, surgery may be necessary to debride infected tissue or to remove damaged bone, with a subsequent need for grafting. Hyperbaric oxygen therapy is sometimes used to reduce the degree of osteoradionecrosis (Marx, Johnson, and Kline, 1985). This therapy facilitates oxygen uptake in blood and related tissues, which in turn improves vascularity, thus reducing tissue damage and related osteoradionecrotic changes following extractions (Marx, 1983a, 1983b).

Once primary medical treatment (surgery, radiotherapy) has been completed, the goals of rehabilitation are to maintain or restore anatomical structure and physiological function. However, if the lesion is extensive and destructive, with distant spread, and subsequent treatment is contraindicated, palliative care may be initiated instead. Palliative care focuses on pain control and maintaining some level of nutrition, respiration, and human dignity.

**Defects of the Maxilla and Velum.** Surgery is the preferred method for treating cancer of the maxilla. Such procedures vary in the extent of resection and may be performed transorally or transfacially. Because surgical extirpation of maxillary tumors may involve extensive resection, a significant reduction in the essential structures for speech may exist post-treatment. Surgical resection of the gum ridge (alveolus) as well as the hard and soft palates may occur. Although defects of the alveolus may be augmented quite well prosthodontically, more significant surgical resections of the hard and soft palate frequently have a dramatic influence on speech production. The primary deficit observed is velopharyngeal incompetence due to structural defects that eliminate the ability to effectively seal the oral cavity from the nasal cavity (Brown et al., 2000).

When considerable portions of both the hard and the soft palate are removed, the integrity of the oral valve for articulatory shaping and the subsequent demands of the resonance system are substantially affected, often with profound influences on oral communication. The goals of rehabilitation include reducing the surgical defect through prosthetic management so that the individual can eat, drink, and attain some level of functional speech. Rehabilitation involves a multidisciplinary team and occurs over several months, during which the postsurgical anatomy changes and the prosthesis is changed in tandem. Healing and other effects of treatment also need to be addressed during this period.

**Defects of the Tongue, Floor of Mouth, Alveolus, Lip, and Mandible.** Surgery for traumatic injury or tumors that invade the mandible, tongue, and floor of the mouth is frequently more radical than surgery for injuries or tumors of the maxilla. The mandible and tongue are movable structures that are intricately involved in swallowing and speech production. Focal excisions may not result in any significant level of noticeable change in deglutition, swallowing, or speech. However, when larger resections are performed on the tongue (partial or total glossectomy) or floor of the mouth, reconstruction with one of a variety of flap procedures is necessary. Large resections and reconstructions also increase the potential for speech disruption because of limited mobility of the reconstructed areas (e.g., portions of the tongue). Defects in the alveolar ridge or lip may have variable effects on speech and general oral function. The literature in this area is scant to nonexistent; however, changes to the lip and alveolus may certainly limit the production of specific speech sounds. For example, anterior sounds in which the lip is an active articulator (e.g., labiodental sounds), sounds that rely on pressure build-up (e.g., stop-approximate sounds), or sounds that require a fixed position for continuous generation or oral airway turbulence (e.g., lingual-alveolar sounds), among others, may be altered by surgical ablation of these structures. In many instances, these types of problems can be addressed using methods commonly employed for articulation therapy. Slowing the speech rate may help to augment articulatory precision in the presence of such defects following surgical treatment.

Isolated mandibular tumors are rare, but regional spread of cancer to the mandible from other oral structures is not uncommon. Extended lesions involving the mandible often require surgical resection. Because cancer has invaded bone, resections are often substan-
tial, subsequently requiring reconstruction with donor bone or plate reconstruction methods. However, changes in the mobility of the mandible pose a risk for changes in the overall acoustic structure of speech due to changes in oral resonance. Prosthetic rehabilitation for jaw defects resulting from injury or malignancy is essential (Adisman, 1990). However, mandibular reconstructions may permit the individual to manipulate the mandible with relative efficiency for speech and swallowing movements. Again, this area is not well documented in the clinical literature.

**Oral Articulatory Evaluation.** Total or partial removal of the tongue (glossectomy) or resection of the anterior maxilla or mandible results in significant speech impairment. The results of peripheral structural ablations of the oral speech system are best addressed using the methods recommended for treatment of the dysarthrias (e.g., Kent et al., 1989). Acoustic analysis, including evaluation of formant frequency, amplitude, and spectral moments, is of clinical value (Gillis and Leonard, 1982; Leonard and Gillis, 1982; Tobey and Lineks, 1989). Acoustic considerations relative to speech intelligibility should also be addressed in this population using tools developed for the dysarthrias (Kent et al., 1989).

**Prosthetic Management.** To optimize the care of each patient, the expertise of multiple professionals, including a dentist, prosthodontist, head and neck surgeon, prosthodontist, speech-language pathologist, and radiologist, is necessary. Surgical treatment of maxillary tumors can result in a variety of problems postsurgically. For example, leakage of food or drink into the nasal cavity significantly disrupts eating and drinking. This may be compounded by difficulties in chewing food or by dysphagia, so that nutritional problems must always be considered. Additionally, such surgical defects reduce articulatory precision and change the acoustic structure of speech because of changes in the volume of the vocal tract as well as resultant hypernasality and nasal emission (abnormal and perhaps excessive flow of air through the nasal cavity). Together, these types of changes to the peripheral oral structures will almost certainly result in decreased speech intelligibility and communication. In order for speech improvement to occur, the surgical defect must be augmented with a prosthesis. This is typically a multistage process, with the first obturator being placed in situ at the time of surgery. The initial obturator is maintained for 1–2 weeks, at which time an interim device is fabricated. Most individuals use the interim device until complete healing has occurred and related treatments are completed, usually anywhere from 3 to 6 months after surgery. At this time a final or definitive obturator is fabricated. Although some minor adjustments to the definitive prosthesis are not uncommon, this obturator will be maintained permanently (Desjardins, 1977, 1978; Anderson et al., 1992).

**Prosthetic Management of Surgical Defects.** Early oral rehabilitation is essential, for reasons that go beyond speech concerns. The surgical obturator allows careful packing of the wound site, keeps food and other debris from entering the wound, permits eating, drinking, and swallowing without a nasogastric tube, and allows immediate oral communication (Doyle, 1999; Leeper and Gratton, 1999). However, certain factors may alter the course of rehabilitation. Fabrication of a prosthesis is more difficult when teeth are absent, as this creates problems in fitting and retention of the prosthesis, which may result in secondary problems. When the jaw is disrupted, changes in symmetry may emerge that may influence chewing and swallowing. For soft palate defects, early prostheses can be modified throughout the course of treatment to facilitate swallowing and speech. Because respiration is paramount, the devices must be created so that they do not impede normal breathing.

In general, intraoral prostheses for those with surgical defects of peripheral oral structures seek to maintain the oral and nasal cavities as separate entities and to reduce velopharyngeal orifice area. Reduction of surgically induced velopharyngeal incompetence is essential to enhancing residual speech capabilities, particularly following larger resections of the maxilla. In addition to the obvious defects in the structural integrity of the peripheral oral mechanism, treatment may disrupt neural processes, including both afferent and efferent components of the system. As such, speech assessment must evaluate the motor capacity as well as the capacity of the sensory system.

When reductions in the ability of the tongue to approximate superior structures of the oral cavity exist after treatment, attempts to facilitate this ability may be achieved by constructing a “palatal drop” prosthesis, which supplements lingual inability by bringing the new hard palate inferior in the oral cavity. The individual may then have greater ability to make the necessary oral contacts to improve articulatory precision and speech intelligibility (e.g., Aramany et al., 1982; Gillis and Leonard, 1983). Such a prosthesis is useful in patients who have healed after undergoing a maxillectomy or maxillary-mandibular resection. Prosthetic adaptations of this type may also benefit eating and swallowing. These devices serve both speech and swallowing by helping to reduce velopharyngeal defects.

**Speech Assessment**

Systematic assessment of the peripheral speech mechanism includes formal evaluation of all subsystems—respiratory, laryngeal, velopharyngeal, and articulatory. As each component of the system is assessed, its relative contributions to the overall speech deficits observed can be better defined for purposes of treatment monitoring. Subsystem evaluation may entail instrumentally based assessments (speech aerodynamics and acoustics), which elicit information on the aeromechanical relationship to oral port size and tongue–hard palate valving (Warren and DuBois, 1964), auditory-perceptual evaluations of speech or voice parameters, and measures of speech intelligibility (Leeper, Sills, and Charles, 1993; Leeper and...
Gratton, 1999). Use of Netsell and Daniel’s (1979) physiological model allows the clinician to identify the relative contribution of each speech subsystem and to target appropriate therapy techniques in an effort to optimize the system. The physiologic approach also permits continuous evaluation of each component in a comparative manner for prosthesis-in and prosthesis-out conditions. Such evaluations aid in fine-tuning prosthetic management.

Once obturation occurs, direct speech treatment goals may include improving the control of the respiratory support system for speech, determining and maintaining the optimal speech rate, increasing the accuracy of specific sound production (or potentially the directed compensation of sound substitutions), improving intelligibility through overarticulation, modulating vocal intensity to improve oral resonance, and related tasks that seek to improve intelligibility. Treatment goals may best be achieved by using the methods previously described for the dysarthrias (Dworkin, 1991).

—Philip C. Doyle

References


Further Readings


Psychogenic Voice Disorders: Direct Therapy

Psychogenic voice disorder refers to a voice disorder that is a manifestation of some confirmed psychological disequilibrium (Aronson, 1990). In its purest form, the psychogenic voice disorder is not associated with structural laryngeal changes or frank central or peripheral nervous system pathology. It is asserted that the larynx, by virtue of its neural connections to emotional centers within the brain, is vulnerable to excess or poorly regulated musculoskeletal tension arising from stress, conflict, fear, and emotional inhibition (Case, 1996). Such dysregulated laryngeal muscle tension can interfere with normal voice and give rise to complete aphonia (i.e., whispered speech) or partial voice loss (dysphonia). Although numerous theories have been offered to explain psychogenic voice loss, the precise mechanisms underlying such psychologically based disorders have not been fully elucidated (see Functional Voice Disorders for a review). Despite considerable controversy surrounding causal mechanisms, the clinical voice literature is replete with evidence that symptomatic voice therapy for psychogenic disorders can often result in rapid and dramatic voice improvement (Koufman and Blalock, 1982; Aronson, 1990; Milutinovic, 1990; Carding and Horsley, 1992; Roy and Leeper, 1993; Gunther et al., 1996; Roy et al., 1997; Andersson and Schalen, 1998; Carding, Horsley, and Docherty, 1999; Stemple, 2000). The following discussion considers voice therapy techniques aimed at directly alleviating vocal symptoms without specific attention to the putative psychological dysfunction underlying the disorder.

Before symptomatic therapy is begun, the laryngological findings are reviewed with the patient, and he or she is reassured regarding the absence of any structural laryngeal pathology. An explanation of the problem is then provided by the clinician. While the specific approach and emphasis vary among clinicians, the discussion typically includes some explanation of the unoward effects of excess or dysregulated muscle tension on voice production and its probable link to stress, situational conflicts, or other psychological precursors that were identified during the interview. The confident clinician provides brief information regarding the therapy plan and the likelihood of a positive outcome.

Because excess or dysregulated laryngeal muscle tension is frequently offered as the proximal cause of the psychogenic voice disorder, many voice therapies including yawn-sigh resonant voice therapy, visual and electromyographic biofeedback, chewing therapy, progressive relaxation, and circumlaryngeal massage aimed at reducing or rebalancing such tension (Boone and McFarlane, 2000). Prolonged hypercontraction of laryngeal muscles is often associated with elevation of the larynx and hyoid bone, with associated pain and discomfort when the circumlaryngeal region is palpated. Aronson (1990) and Roy and Bress (1998) have described manual techniques to determine the presence and degree of laryngeal musculoskeletal tension, as well as methods to relieve such tension during the diagnostic assessment and management session. Circumlaryngeal massage is one such treatment approach. Skillfully applied, systematic kneading of the extralaryngeal region is believed to stretch muscle tissue and fascia, promote local circulation with removal of metabolic wastes, relax tense muscles, and relieve pain and discomfort associated with muscle spasms. The hypothesized physical effect of such massage is reduced laryngeal height and stiffness and increased mobility. Once the larynx is “released” and range of motion is normalized, proportional improvement in voice is said to follow. Improvement in voice and reductions in pain and laryngeal height suggest a relief of tension (Roy and Ferguson, 2001). In a similar vein, Roy and Bress (1998) also recently described a number of manual laryngeal repositioning techniques that can stimulate improved voice...
and briefly interrupt patterns of muscle misuse. These brief moments of voice improvement associated with laryngeal reposturing maneuvers are immediately identified for the patient and reinforced. Digital cues can then be faded and the patient taught to rely on sensory feedback (auditory, kinesthetic, and proprioceptive) to maintain improved laryngeal positioning and muscle balance. Any partial relapses or return of abnormal voice during the therapy process can be dealt with by reassurance, verbal reinstruction, or manual cueing. Once the larynx is correctly positioned, recovery of normal voice can occur rapidly.

Certain patients with psychogenic dysphonia and aphonia appear to have lost kinesthetic awareness and volitional control over voice production for speech and communication purposes, yet display normal voicing for vegetative or nonspeech acts. For instance, some aphonic and severely dysphonic patients may be able to clear the throat, grunt, cough, sigh, gargle, laugh, hum, or produce a high-pitched squeak with normal or near-normal voice quality. Such preserved voicing for non-speech purposes represents a clue to the capacity for normal phonation. In symptomatic therapy, then, the patient is asked to produce such vocal behaviors. The goal of these voice maneuvers is to elicit even a brief trace of clearer voice so that it may be shaped toward normal quality or extended to longer utterances. These efforts follow a trial-and-error pattern and require the seasoned clinician to be constantly vigilant, listening for any brief moments of clearer voice. When improved voice is elicited, it is instantly reinforced, and the clinician provides immediate feedback regarding the positive change. During this process, the patient needs to be an active participant and is encouraged to continually self-monitor the type and manner of voice produced. Once this brief but relatively normal voice is reliably achieved, it is shaped and extended into sustained vowels, words, simple phrases, and oral reading. When this phase of intervention is successful, the patient is then engaged in casual conversation that begins with basic biographical information and proceeds to brief narratives, and then conversation about any topic and with anyone in the clinical setting. If established, the restored voice is usually maintained without compensatory effort and may improve further during conversation. Finally, the clinician should debrief the patient regarding the cause of the voice improvement, discuss the patient’s feelings about the improved voice, and review possible causes of the problem and the prognosis for maintaining normal voice.

Certainly, direct symptomatic therapy for psychogenic voice disorders can produce rapid changes; however, in some cases voice therapy can be a frustrating and protracted experience for both clinician and patient (Bridger and Epstein, 1983; Fex et al., 1994). The rate of improvement during therapy for psychogenic voice disorders varies. Patients may progress gradually through various stages of dysphonia on their way to normal voice recovery. Other patients will appear to experience sudden return of voice without necessarily transitioning through phases of decreasing severity (Aronson, 1990).

Because there are few studies directly comparing the effectiveness of specific therapy techniques, not much is known about whether one therapy approach for psychogenic voice disorder is superior to another. Although signs of improvement should typically be observed within the first session, some patients may need an extended, intensive treatment session or several sessions, depending on several variables, including the therapy techniques selected, clinician experience and confidence in administering the approach, and patient motivation and tolerance, to mention only a few.

The anecdotal clinical literature suggests that the prognosis for sustained removal of abnormal symptoms in psychogenic aphony or dysphonia may depend on several factors. First, the time between the onset of voice problem and the initiation of therapy may be important. The sooner voice therapy is initiated following the onset of the voice problem, the better the prognosis. If months or years have elapsed, it may be more difficult to eliminate the abnormal symptoms. Second, the more extreme the voice symptoms, the better the prognosis for improvement. Aphony and extreme tension, according to some authorities, may be easier to modify than intermittent or mild dysphonia. Third, if significant secondary gain is present, this may interfere with progress and contribute to a poorer treatment outcome. Finally, if the underlying psychological triggers are no longer active, then normal voice should be established quickly and improvement should be sustained (Aronson, 1990; Duffy, 1995; Case, 1996; Colton and Casper, 1996). As a caveat, however, the foregoing observations have rarely been studied in any objective manner; therefore they are best regarded as clinical impressions rather than factual statements.

The long-term effectiveness of direct voice therapy for psychogenic voice disorders also has not been rigorously evaluated (Pannbacker, 1998; Ramig and Verdolini, 1998). Most clinicians report that relapse is infrequent and isolated, yet others report more frequent post-treatment recurrences. Of the few investigations that exist, the results regarding the durability of voice improvement following direct therapy are mixed (Gunther et al., 1996; Roy et al., 1997; Andersson and Schalen, 1998; Carding, Horsley, and Docherty, 1999). It should be acknowledged that following direct voice therapy, only the symptom of psychological disturbance has been removed, not the disturbance itself (Brodnitz, 1962; Kinzl, Biebl, and Rauchegger, 1988). Therefore, the nature of psychological dysfunction needs to be better understood. If the situational, emotional, or personality features that contributed to the development of the psychogenic voice disorder remain unchanged following behavioral treatment, it would be logical to expect that such persistent factors would increase the probability of future recurrences (Nichol, Morrison, and Rammage, 1993; Andersson and Schalen, 1998). Therefore, in some cases, post-treatment referral to a psychiatrist or psychologist may be necessary to achieve more enduring improvements in the patient’s emotional adjustment and voice function (Butcher, Elias, and Raven, 1993; Roy et al., 1997).
In summary, psychogenic voice disorders are powerful examples of the intimate connection between mind and body. These voice disorders, which occur in the absence of structural laryngeal pathology, often represent some of the most severely disturbed voices encountered by voice pathologists. In an experienced clinician’s hands, direct symptomatic therapy for psychogenic voice disorders can produce rapid and remarkable restoration of normal voice. Much remains to be learned regarding the underlying bases of these disorders and the long-term effect of direct therapeutic interventions.

See also FUNCTIONAL VOICE DISORDERS.

—Nelson Roy

References


The Singing Voice

The functioning of the voice organ in singing is similar to that in speech. Thus, the origin of the sound is the voice source—the pulsating airflow through the glottis. The voice source is mainly controlled by three physiological factors, subglottal pressure, length and stiffness of the vocal folds, and the degree of glottal adduction. These control parameters determine vocal loudness, F0, and mode of phonation, respectively. The voice source is a complex tone composed of a series of harmonic partials of amplitudes decreasing by about 12 dB per octave as measured in flow units. It propagates through the vocal tract and is thereby filtered in a manner determined by its resonance or formant frequencies. These frequencies are determined by the vocal tract shape. For most vowel sounds, the two lowest formant frequencies determine vowel quality, while the higher formant frequencies belong to the personal voice characteristics.

Breathing. Subglottal pressure determines vocal loudness and is therefore used for expressive purposes in singing. It is also varied with F0, such that higher pitches
are sung with higher subglottal pressures than lower pitches. As a consequence, singers need to vary subglottal pressure constantly, adapting it to both loudness and pitch. This is in sharp contrast to speech, where subglottal pressure is much more constant. Singers therefore need to develop a virtuosic control of the breathing apparatus. In addition, subglottal pressures in singing are varied over a larger range than in speech. Thus, while in loud speech subglottal pressure may be raised to 1.5 or 2 kPa, singers may use pressures as high as 4 or 6 kPa for loud tones sung at high pitches.

Subglottal pressure is determined by active forces produced by the breathing muscles and passive forces produced by gravity and the elasticity of the breathing apparatus. Elasticity, generated by the lungs and the rib cage, varies with lung volume. At high lung volumes, elasticity produces an exhalatory force that may amount to 3 kPa or more. At low lung volumes, elasticity contributes an inhalatory force. Whereas in conversational speech, no more than about 15%–20% of the total lung capacity is used, classically trained singers use an average lung volume range that is more than twice as large and occasionally may vary from 100% to 0% of the total vital capacity in long phrases.

As the elasticity forces change from exhalatory at high lung volumes to inhalatory at low lung volumes, they reach an equilibrium at a certain lung volume. This lung volume is called the functional residual capacity (FRC). In tidal breathing, inhalations are started from FRC. In both speech and singing, lung volumes above FRC are preferred. Because much higher lung volumes are used in singing than in speech, singers need to deal with much greater exhalatory elasticity forces.

**Voice Source.** The airflow waveform of the voice source is characterized by quasi-triangular pulses, produced when the vocal folds open the glottis, and followed by horizontal portions near or at zero airflow, produced when the folds close the glottis more or less completely (Fig. 1).

The acoustic significance of the waveform is straightforward. The slope of the source spectrum is determined mainly by the negative peak of the differentiated flow waveform, frequently referred to as the maximum flow declination rate. It represents the main excitation of the vocal tract. This steepness is linearly related to the subglottal pressure in such a way that a doubling of subglottal pressure causes an SPL increase of about 10 dB. The amplitude gain of higher partials is greater than that of lower partials. Thus, if the sound level of a vowel sound is increased by 10 dB, the partials near 3 kHz typically increase by about 17 dB.

The air volume contained in a flow pulse is decisive to the amplitude of the source spectrum fundamental and is strongly influenced by the overall glottal adduction force. Thus, for a given subglottal pressure, a firmer adduction produces a smaller air volume in a pulse, which reduces the amplitude of the fundamental. An exaggerated glottal adduction thus attenuates the fundamental. This phonation mode is generally referred to as hyperfunctional or pressed. The opposite extreme—that is, the habitual use of too faint adduction—is called hypofunctional and prevents the vocal folds from closing the glottis also during the vibratory cycle. As a result, airflow escapes the glottis during the quasi-closed phase. This generates noise and produces a strong fundamental. This phonation mode is often referred to as breathy.

In classical singing, pressed phonation is typically avoided. Instead, singers seem to strive to reduce glottal adduction to the minimum that will still result in glottal closure during the closed phase. This generates a source spectrum with strong high partials and a strong fundamental. This type of phonation has been called flow phonation or resonant voice. In nonclassical singing, on the other hand, pressed phonation is occasionally used for high, loud tones, apparently for expressive purposes.

A main characteristic of classical singing is the vibrato. It corresponds to a quasi-periodic modulation of F0 (Fig. 2). The pitch perceived from a vibrato tone corresponds to its average F0. The modulation frequency, mostly between 5 and 7 Hz, is generally referred to as the vibrato rate and is rather constant for a singer. Curiously enough, however, it tends to increase somewhat toward the end of tones. The peak-to-peak modulation range is varied between nil and less than two semitones, or $F0 \cdot 2^{1/6}$. With increasing age, singers’ vibrato rates tend to decrease by about one-half hertz per decade of years, and vibrato extent tends to increase by about 15 cent per decade.

The vibrato is generated by a rhythmic pulsation of the cricothyroid muscles. When contracting, these muscles cause a stretching of the vocal folds, and so raise F0. The neural origin of this pulsation is not understood. One possibility is that it emanates from a cocontraction of the cricothyroid and vocalis muscles.

In speech, pitch is perceived in a continuous fashion, such that a continuous variation in F0 is heard as a continuous variation of pitch. In music, on the other hand, pitch is perceived categorically, where the categories are scale tones or the intervals between them. Thus,
the F0 continuum is divided logarithmically into a series of bins, each of which corresponds to a scale tone. The width of each scale tone is approximately 6% wide, and the center frequency of a scale tone is $2^{1/12}$ higher than its lower neighbor.

The demands for pitch accuracy are quite high in singing. Experts generally find that a tone is out of tune when it deviates from the target F0 by more than about 7 cent, or 0.07 of a semitone interval. This corresponds to less than one-tenth of a typical vibrato extent. The target F0 generally agrees with equal-tempered tuning, where the interval between adjacent scale tones corresponds to the F0 ratio of $1:2^{1/12}$. However, apparently depending on the musical context, the target F0 for a scale tone may deviate from its value in equal-tempered tuning by about a tenth of a semitone interval.

Resonance. The formant frequencies in classical singing differ between voice classifications. Thus, basses have lower formant frequencies than baritones, who have lower formant frequencies than tenors. These differences probably reflect differences in vocal tract length. The formant frequencies also deviate from those typically found in speech. For example, the second formant of the vowel [i] is generally considerably lower in classical singing than in speech, such that the vowel quality approaches that of the vowel [y].

These formant frequency deviations are related to the singer’s formant, a marked spectrum envelope peak between approximately 2.5 and 3 kHz. If this resonance is appropriately tuned, it will provide a formant cluster.

A common method to achieve a wide pharynx seems to be to lower the larynx, which is typically observed in classically trained singers. Lowering the larynx lengthens the pharynx cavity. As F2 of the vowel [i] is mainly dependent on the pharynx length, it will be lowered by a lowering of the larynx. In nonclassical singing, more speechlike formant frequencies are used, and no singer’s formant is produced.

The center frequency of the singer’s formant varies between voice classifications. On average, it tends to be about 2.4, 2.6, and 2.8 kHz for basses, baritones, and tenors, respectively. These differences, which contribute significantly to the characteristic voice qualities of these classifications, probably reflect differences in vocal tract length.

The center frequency of the singer’s formant varies between voice classifications. On average, it tends to be about 2.4, 2.6, and 2.8 kHz for basses, baritones, and tenors, respectively. These differences, which contribute significantly to the characteristic voice qualities of these classifications, probably reflect differences in vocal tract length.

The singer’s formant spectrum peak is particularly prominent in bass and baritone singers. In tenors and altos it is less prominent and in sopranos it is generally not observable. Thus, sopranos do not seem to produce a singer’s formant.

The approximate pitch range of a singer is about two octaves. Typical ranges for basses, baritones, tenors,
altos, and sopranos are E2–E4 (82–330 Hz), G2–G4 (98–392 Hz), C3–C5 (131–523 Hz), F3–F5 (175–698 Hz), and C4–C6 (262–1047 Hz), respectively. This implies that F0 is often higher than the typical value of F1 in some vowels. Singers, however, seem to avoid the situation in which F0 is higher than F1. Instead, they increase F1 so that it is always higher than F0. For the vowel [a], this is achieved by widening the jaw opening: the higher the pitch, the wider the jaw opening. For other vowels, singers seem first to reduce the tongue constriction of the vocal tract, and resort to a widening of the jaw opening when the effect of this neutralization of the articulation fails to produce further increase of F1.

Because F1 and F2 are decisive to the perception of vowels, the substantial departures from the typical formant frequency values in speech affect vowel identification. Yet vowel identification is surprisingly successful also at high F0. Most isolated vowel sounds can be correctly identified up to an F0 of about 500 Hz. Above this frequency, identification deteriorates quickly and remains low for most vowels at F0 higher than 700 Hz. In reality, however, text intelligibility can be greatly improved by consonants.

Health Risks. Because singers are extremely dependent on the perfect functioning of their voices, they often need medical attention. A frequent origin of their voice disorders is a cold, which typically causes dryness of the glottal mucosa. This disturbs the normal function of the vocal folds. Also relevant would be their use of high subglottal pressures. An inappropriate vocal technique, sometimes associated with a habitually exaggerated glottal adduction or with singing in a too high pitch range, also tends to cause voice disorders, which in some cases may lead to developing vocal nodules. Such nodules generally disappear after voice rest, and surgical treatment is mostly considered inappropriate.

See also VOICE ACOUSTICS.

—Johan Sundberg

References


Vocal Hygiene

Vocal hygiene has been part of the voice treatment literature continuously since the publication of Mackenzie’s The Hygiene of Vocal Organs, in 1886. In it the author, a noted otolaryngologist, described many magical prescriptions used by famous singers to care for their voices. In 1911, a German work by Barth included a chapter with detailed discussion of vocal hygiene. The ideas about vocal hygiene expressed in this book were similar to those expressed in the current literature. Concern was raised about the effects of tobacco, alcohol, loud and excessive talking, hormones, faulty habits, and diet on the voice. Another classic text was Diseases and Injuries of the Larynx, published in 1942. The authors, Jackson and Jackson, implicated various vocal abuses as the primary causes of voice disorders, and cited rest and refrain from the behavior as the appropriate treatment. Luchsinger and Arnold (1965) stressed the need for attention to the physiological norm as the primary postulate of vocal hygiene and preventive laryngeal medicine. Remarkably, these authors discussed the importance of this type of attention not only for teachers and voice professionals, but also for children in the classroom. Subsequently, virtually all voice texts have addressed the issue of vocal hygiene.

Both the general public and professionals in numerous disciplines commonly use the term hygiene. The 29th edition of Dorland’s Medical Dictionary defines it as “the science of health and its preservation.” Thus, we can take vocal hygiene to mean the science of vocal health and the proper care of the vocal mechanism. Despite long-held beliefs about the value of certain activities most frequently discussed as constituting vocal hygiene, the science on which these ideas are based was, until quite recently, more implied and deduced than specific.

Patient education and vocal hygiene are both integral to voice therapy. Persons who are educated about the structure and function of the phonatory mechanism are better able to grasp the need for care to restore it to health and to maintain its health. Thus, the goal of patient education is understanding. Vocal hygiene, on the other hand, focuses on changing an individual's vocal behavior. In some instances, a therapy program may be based completely on vocal hygiene. More frequently, however, vocal hygiene is but one spoke in a total therapy program that also includes directed instruction in voice production techniques.

Although there are commonalities among vocal hygiene programs regardless of the pathophysiology of the voice disorder, that pathophysiology should dictate some specific differences in the vocal hygiene approach. In addition to the nature of the voice disorder, factors
such as timing of the program relative to surgery (i.e., pre or post), and whether the vocal hygiene training stands alone or is but one aspect of a more extensive therapy process, must also inform specific aspects of the vocal hygiene program.

Hydration and environmental humidification are particularly important to the health of the voice, and, as such, should be a focus of all vocal hygiene programs. A number of authors have studied the effects of hydration and dehydration of vocal fold mucosa and viscosity of the folds on phonation threshold pressure (PTP). (PTP is the minimum subglottal pressure required to initiate and maintain vocal fold vibration.) For example, Verdolini et al. (1990, 1994) studied PTP in normal speakers subjected to hydrating and dehydrating conditions. Both PTP and self-perceived vocal effort were lower after hydration. Jiang, Ng, and Hanson (1999) showed that vocal fold oscillations cease in a matter of minutes in fresh excised canine larynges deprived of humidified air. Rehydration by dripping saline onto the folds restored the oscillations, demonstrating the need for hydration and surface moisture for lower PTP.

In one light, viscosity is a measure of the resistance to deformation of the vocal fold tissue. Viscosity is increased by hydration and decreased by drying—hence the importance of vocal fold hydration to ease of phonation. Moreover, it appears that the body has robust cellular and neurophysiological mechanisms to conserve the necessary hydration of airway tissues. In a study of patients undergoing dialysis, rapid removal of significant amounts of body water increased PTP and was associated with symptoms of mild vocal dysfunction in some patients. Restoration of the body fluid reversed this trend (Fisher et al., 2001). Jiang, Lin, and Hanson (2000) noted the presence of mucous glands in the tissue of the vestibular folds and observed that these glands distribute a very important layer of lubricating mucus to the surface of the vocal folds. Environmental hydration facilitates the vocal fold vibratory behavior, mainly because of the increased water content in this mucous layer and in the superficial epithelium. The viscosity of secretions is thickened with ingestion of foods or medications with a drying or diuretic effect, radiation therapy, inadequate fluid intake, and the reduction in mucus production in aging.

Thus, there appear to be a number of mechanisms, not yet fully understood, by which the hydration of vocal fold mucosa and the viscosity of the vocal folds are directly involved in the effort required to initiate and maintain phonation. Both environmental humidity and surface hydration are important physiological factors in determining the energy needed to sustain phonation. External or superficial hydration may occur as a by-product of drinking large amounts of water, which increases the secretions in and around the larynx and lowers the viscosity of those secretions. Steam inhalation and environmental humidification further hydrate the surface of the vocal folds, and mucolytic agents may decrease the viscosity of the vocal folds. Clearly, the lower the phonation threshold pressure, the less air pressure is required and the greater is the ease of phonating. Many questions remain in this area, such as the most effective method of hydration.

Other major components of vocal hygiene programs are reducing vocal intensity by eliminating shouting or speaking above high ambient noise levels, avoiding frequent throat clearing and other phonotraumatic behaviors. The force of collision (or impact) of the vocal folds has been described by Titze (1994) as proportional to vibrational amplitude and vibrational frequency. This was explored further in phonation by Jiang and Titze (1994), who showed that intraglottal contact increases with increased vocal fold adduction. Titze (1994) theorized that if a vibrational dose reaches and exceeds a threshold level in a predisposed individual, tissue injury will probably ensue. This lends support to the widespread belief that loud and excessive voice use, and indeed other forms of harsh vocal productions, can cause vocal fold pathology. It also supports the view that teachers and others in vocally demanding professions are prone to vibration overdose, with inadequate recovery time. Thus, the stage is set for cyclic tissue injury, repair, and eventual voice or tissue change.

The complexity of vocal physiology suggests a direct connection between viscosity and hydration, phonation threshold pressure, and the effects of collision and shearing forces. The greater the viscosity of vocal fold tissue, the higher the PTP that is required and the greater is the internal friction or shearing force in the vocal fold. These effects may explain vocal fold injuries, particularly with long-term vocal use that involves increased impact stress on the tissues during collision and shearing stresses (Jiang and Titze, 1994). Thus, issues of collision and the impact forces associated with increased loudness and phonotraumatic vocalization are appropriately addressed in vocal hygiene programs and in directed therapy approaches.

Reflux, both gastroesophageal and laryngopharyngeal, affects the health of the larynx and pharynx. Gastric acid and gastric pepsin, the latter implicated in the delayed healing of submucosal laryngeal injury (Koufman, 1991), have been found in refluxed material. Laryngopharyngeal reflux has been implicated in a long list of laryngeal conditions, including chronic or intermittent dysphonia, vocal fatigue, chronic throat clearing, reflux laryngitis, vocal nodules, and malignant tissue changes. Treatment may include dietary changes, lifestyle modifications, and medication. Surgery is usually a treatment of last resort. Caffeine, tobacco, alcohol, fried foods, and excessive food intake have all been implicated in exacerbating the symptoms of laryngopharyngeal reflux. Thus, vocal hygiene programs that address healthy diet and lifestyle and that include reflux precautions appear to be well-founded. It is now common practice for patients scheduled for any laryngeal surgery to be placed on a preoperative course of antireflux medication that will be continued through the postoperative healing stage. Although this is clearly a medical treatment, the speech-language pathologist should provide information and supportive guidance through vocal hygiene instruction to ensure that patients follow the prescribed protocol.
An unanswered question is whether a vocal hygiene therapy program alone is adequate treatment for vocal problems. Roy et al. (2001) found no significant improvement in a group of teachers with voice disorders after a course of didactic training in vocal hygiene. Teachers who received a directed voice therapy program (Vocal Function Exercises), however, experienced significant improvement. It should be noted that the vocal hygiene program used in this study, being purely didactic and requiring no activity on the part of the participants, might more appropriately be described as a patient education program. Chan (1994) reported that a group of non-voice-disordered kindergarten teachers did show positive behavioral changes following a program of vocal education and hygiene. In another study, Roy et al. (2002) examined the outcome of voice amplification versus vocal hygiene instruction in a group of voice-disordered teachers. Most pairwise contrasts directly comparing the effects of the two approaches failed to reach significance. Although the vocal hygiene group showed changes in the desired direction on all dependent measures, the study results suggest that the benefits of amplification may have exceeded those of vocal hygiene instruction. Of note, the amplification group reported higher levels of extracurricular compliance with the program than the vocal hygiene group. This bears out the received wisdom that it is easier to take a pill—or wear an amplification device—than to change habits.

Although study results are mixed, there is insufficient evidence to suggest that vocal hygiene instruction be abandoned. The underlying rationale for vocal hygiene is sufficiently compelling that a vocal hygiene program should continue to be a component of a broad-based voice therapy intervention.

—Janina K. Casper

References


Vocal Production System: Evolution

The human vocal production system is similar in broad outline to that of other terrestrial vertebrates. All tetrapods (nonfish vertebrates: amphibians, reptiles, birds, and mammals) inherit from a common ancestor three key components: (1) a respiratory system with lungs; (2) a larynx that acts primarily as a quick-closing gate to protect the lungs, and often secondarily to produce sound; and (3) a supralaryngeal vocal tract which filters this sound before emitting it into the environment. Despite this shared plan, a wide variety of interesting modifications of the vocal production system are known. The functioning of the basic tetrapod vocal production system can be understood within the theoretical framework of the myoelastic-aerodynamic and source/filter theories familiar to speech scientists.

The lungs and attendant respiratory musculature provide the air stream powering phonation. In primitive air-breathing vertebrates, the lungs were inflated by rhythmic compression of the oral cavity, or “buccal pumping,” and this system is still used by lungfish and amphibians (Brainerd and Ditelberg, 1993). Inspiration by active expansion of the thorax evolved later, in the ancestor of reptiles, birds, and mammals. This was powered originally by the intercostal muscles (as in lizards or crocodilians) and later (in mammals only) by a muscular diaphragm (Liem, 1985). Phonation is typically powered by passive deflation of the elastic lungs, or in some cases by active compression of the hypaxial musculature. In many frogs, air expired from the lungs during phonation is captured in an elastic air sac, which then deflates, returning the air to the lungs. This allows frogs to produce multiple calls from the same volume.
of air. The inflated sac also increases the efficiency with which sound is radiated into the environment (Gans, 1973).

The lungs are protected by a larynx in all tetrapods. This structure primitively includes a pair of barlike cartilages that can be separated (for breathing) or pushed together (to seal the airway) (Negus, 1949). Expiration through the partially closed larynx creates a turbulent hiss—perhaps the most primitive vocalization, which virtually all tetrapods can produce. However, more sophisticated vocalizations became possible after the innovation of elastic membranes within the larynx, the vocal cords, which are found in most frogs, vocal reptiles (geckos, crocodilians), and mammals. Although the larynx in these species can support a wide variety of vocalizations, its primary function as a protective gateway appears to have constrained laryngeal anatomy. In birds, a novel phonatory structure called the syrinx evolved at the base of the trachea. Dedicated to vocal production, and freed from the necessity of tracheal protection, the avian syrinx is a remarkably diverse structure undergoing the great variety of bird sounds (King, 1989).

Although our knowledge of animal phonation is still limited, phonation in nonhumans appears to follow the principles of the myoelastic-aerodynamic theory of human phonation. The airflow from the lungs sets the vocal folds (or syringeal membranes) into vibration, and the rate of vibration is passively determined by the size and tension of these tissues. Vibration at a particular frequency does not typically require neural activity at that frequency. Thus, relatively normal phonation can be obtained by blowing moist air through an excised larynx, and rodents and bats can produce ultrasonic vocalizations at 40 kHz and higher (Suthers and Fattu, 1973). However, cat purring relies on an active tensing of the vocal fold musculature at the 20–30 Hz fundamental frequency of the purr (Frazer Sissom, Rice, and Peters, 1991). During phonation, the movements of the vocal folds can be periodic and stable (leading to tonal sounds) or highly aperiodic or even chaotic (e.g., in screams); while such aperiodic vocalizations are rare in nonpathological human voices, they can be important in animal vocal repertoires (Fitch, Neubauer, and Herzel, 2000).

Because the length of the vocal folds determines the lowest frequency at which they could vibrate (Titze, 1994), with long folds producing lower frequencies, one might expect that a low fundamental would provide a reliable indication of large body size. However, the size of the larynx is not tightly constrained by body size. Thus, a huge larynx has independently evolved in many mammal species, probably in response to selection for low-pitched voices (Fig. 1.4, B). For example, in howler monkeys (genus *Alouatta*) the larynx and hyoid have grown to fill the space between mandible and sternum, giving these small monkeys remarkably impressive and low-pitched voices (Kelemen and Sade, 1960). The most extreme example of laryngeal hypertrophy is seen in the hammerhead bat *Hypsignathus monstrosus*, in which the larynx of males expands to fill the entire thoracic cavity, pushing the heart, lungs, and trachea down into the abdomen (Schneider, Kuhn, and Keleman, 1967). A similar though less impressive increase in larynx dimensions is observed in human males and is partially responsible for the voice change at puberty (Titze, 1989).

Sounds created by the larynx must pass through the air contained in the pharyngeal, oral, and nasal cavities, collectively termed the supralaryngeal vocal tract or simply vocal tract. Like any column of air, this air has mass and elasticity and vibrates preferentially at certain resonant frequencies. Vocal tract resonances are termed formants (from the Latin formare, to shape): they act as filters to shape the spectrum of the vocal output. Because all tetrapods have a vocal tract, all have formants. Formant frequencies are determined by the length and shape of the vocal tract. Because the vocal tract in mammals rests within the confines of the head, and skull size and body size are tightly linked (Fitch, 2000b), formant frequencies provide a possible indicator of body size not as easily “faked” as the laryngeal cue of fundamental frequency. Large animals have long vocal tracts and low formants. Together with demonstrations of formant perception by nonhuman animals (Sommers et al., 1992; Fitch and Kelley, 2000), this suggests that formants may have provided a cue to size in primitive vertebrates (Fitch, 1997). However, it is possible to break the anatomical link between vocal tract length and body size, and some intriguing morphological adaptations have arisen to elongate the vocal tract (presumably resulting from selection to sound larger; Fig. 1C–E). Elongations of the nasal vocal tract are seen in the long nose of male proboscis monkeys or the impressive nasal crests of hadrosaur dinosaurs (Weishampel, 1981). Vocal tract elongation can also be achieved by lowering the larynx; this is seen in extreme form in the red deer *Cervus elaphus*, which retracts the larynx to the sternum during territorial roaring (Fitch and Reby, 2001). Again, a similar change occurs in human males at puberty: the larynx descends slightly to give men a longer vocal tract and lower formants than same-sized women (Fitch and Giedd, 1999).

Human speech is thus produced by the same conservative vocal production system of lungs, larynx, and vocal tract shared by all tetrapods. However, the evolution of the human speech apparatus involved several important changes. One was the loss of laryngeal air sacs. All great apes possess large balloon-like sacs that open into the larynx directly above the glottis (Negus, 1949; Schön Ybarra, 1995). Parsimony suggests that the common ancestor of apes and humans also had such air sacs, which were subsequently lost in human evolution. However, air sacs are occasionally observed in humans in pathological situations, a laryngocele is a congenital or acquired air sac that is attached to the larynx through the laryngeal ventricle at precisely the same location as the human larynx (Schneider, Kuhn, and Keleman, 1967). A similar though less impressive increase in larynx dimensions is observed in human males and is partially responsible for the voice change at puberty (Titze, 1989).

Although the function of air sacs in ape vocalizations is not understood, the significance of their loss in human evolution is unknown.
A second change in the vocal production system during human evolution was the descent of the larynx from its normal mammalian position high in the throat to a lower position in the neck (Negus, 1949). In the 1960s, speech scientists realized that this “descended larynx” allows humans to produce a wider variety of formant patterns than would be possible with a high larynx (Lieberman, Klatt, and Wilson, 1969). In particular, the “point vowels” /i, a, u/ seem to be impossible to attain unless the tongue body is bent and able to move freely within the oropharyngeal cavity. Given the existence of these vowels in virtually all languages (Maddieson, 1984), speech typical of modern humans appears to require a descended larynx. Of course, all mammals can produce a diversity of sounds, which could have served a simpler speech system. Also, most mammals appear to lower the larynx during vocalization (Fitch, 2000a), lessening the gap between humans and other animals. Despite these caveats, the descended larynx is clearly an important component of human spoken language (Lieberman, 1984). The existence of nonhuman mammals with a descended larynx raises the possibility that this trait initially arose to exaggerate size in early hominids and was later coopted for use in speech (Fitch and Reby, 2001). Finally, recent fossils suggest that an expansion of the thoracic intervertebral canal occurred during the evolution of Homo some time after the earliest Homo erectus (MacLarnon and Hewitt, 1999). This change may be associated with an increase in breathing control necessary for singing and speech in our own species.

See also VOCALIZATION, NEURAL MECHANISMS OF.

—W. Tecumseh Fitch

References


Vocalization, Neural Mechanisms of

The capacity for speech and language separates humans from other animals and is the cornerstone of our intellectual and creative abilities. This capacity evolved from rudimentary forms of communication in the ancestors of humans. By studying these mechanisms in animals that represent stages of phylogenetic development, we can gain insight into the neural control of human speech that is necessary for understanding many disorders of human communication.

Vocalization is an integral part of speech and is widespread in mammalian aural communication systems. The limbic system, a group of neural structures controlling motivation and emotion, also controls most mammalian vocalizations. Although there are little supporting empirical data, many human emotional vocalizations probably involve the limbic system. This discussion considers the limbic system and those neural mechanisms thought to be necessary for normal speech and language to occur.

The anterior cingulate gyrus (ACG), which lies on the mesial surface of the frontal cortex just above and anterior to the genu of the corpus callosum, is considered part of the limbic system (Fig. 1). Electrical stimulation of the ACG in monkeys elicits vocalization and autonomic responses (Jürgens, 1994). Monkeys become mute when the ACG is lesioned, and single neurons in the ACG become active with vocalization or in response to vocalizations from conspecifics (Sutton, Larson, and Lindeman, 1974; Müller-Preuss, 1988; West and Larson, 1995). Electrical stimulation of the ACG in humans may also result in oral movements or postural distortions representative of an “archaic” level of behavior (Brown, 1988). Damage to the ACG in humans results in akinetic mutism that is accompanied by open eyes, a fixed gaze, lack of limb movement, lack of apparent affect, and nonreactance to painful stimuli (Jürgens and von Cramon, 1982). These symptoms reflect a lack of drive to initiate vocalization and many other behaviors (Brown, 1988). During recovery, a patient’s ability to communicate gradually returns, first as a whisper, then with vocalization. However, the vocalizations lack prosodic features and are characterized as expressionless (Jürgens and von Cramon, 1982). These observations support the view that the ACG controls motivation for primitive forms of behavior, including prelinguistic vocalization.

The ACG has reciprocal connections with several cortical and subcortical sites, including a premotor area homologous with Broca’s area, the superior temporal gyrus, the posterior cingulate gyrus, and the supplementary motor area (SMA) (Müller-Preuss, Newman, and Jürgens, 1980). Electrical stimulation of the SMA elicits vocalization, speech arrests, hesitation, distortions, and palilalic iterations (Brown, 1988). Damage to the SMA may result in mutism, poor initiation of speech, or repetitive utterances, and during recovery, patients often go through a period in which the production of
propositional speech remains severely impaired but nonpropositional speech (e.g., counting) remains relatively unaffected (Brown, 1988). Studies of other motor systems in primates suggest that the SMA is involved in selection and initiation of a remembered motor act or the correct sequencing of motor acts (Picard and Strick, 1997). Speech and vocalization fall into these categories, as both are remembered and require proper sequencing. Output from the SMA to other vocalization motor areas is a subsequent stage in the execution of vocalization.

The ACG is also connected with the perisylvian cortex of the left hemisphere (Müller-Preuss, Newman, and Jürgens, 1980), an area important for speech and language. Damage to Broca’s area may cause total or partial mutism, along with expressive aphasia or apraxia (Duffy, 1995). However, vocalization is less frequently affected than speech articulation or language (Duffy, 1995), and the effect is usually temporary. In some cases, aphasis may arise from widespread damage, and recovery following therapy may suggest a diffuse, motivational, or psychogenic etiology (Sapir and Aronson, 1987). Mutism seems to occur more frequently when the opercular region of the pre- and postcentral gyri is damaged bilaterally or when the damage extends deep into the cortex, affecting the insula and possibly the basal ganglia (Jürgens, Kirzinger, and von Cramon, 1982; Starkstein, Berthier, and Leiguarda, 1988; Duffy, 1995). Additional evidence linking the insula to vocalization comes from recent studies of apraxia in humans (Dronkers, 1996) and findings of increased blood flow in the insula during singing (Perry et al., 1999). Further research is necessary to determine whether the opercular cortex alone or deeper structures (e.g., insula) are important for vocalization. In specific cases, it is necessary to know whether mutism results from psychogenic or physiological mechanisms.

The perisylvian cortex may control vocalization by one or more pathways to the medulla. The perisylvian cortex is reciprocally connected to the ACG, which projects to midbrain mechanisms involved in vocalization. The perisylvian cortex also projects directly to the medulla, where motor neurons controlling laryngeal muscles are located (Kuypers, 1958). These neuroanatomical projections are supported by observations of a short time delay (13 ms) between stimulation of the cortex and excitation of laryngeal muscles (Ludlow and Lou, 1996). The perisylvian cortex also includes the right superior temporal gyrus and Heschl’s gyrus, which are preferentially active for perception of complex tones, singing, perception of one’s own voice, and perhaps control of the voice by self-monitoring auditory feedback (Perry et al., 1999; Belin et al., 2000).

The other widely studied limbic system structure known for its role in vocalization, is the midbrain periaqueductal gray (PAG) (Jürgens, 1994). Lesions of the PAG in humans and animals lead to mutism (Jürgens, 1994). Electrical and chemical stimulation of the PAG in many animal species elicits species-specific vocalizations (Jürgens, 1994). A variety of techniques have shown that PAG neurons, utilizing excitatory amino acid transmitters (glutamate), activate or suppress coordinated groups of oral, facial, respiratory, and laryngeal muscles for species-specific vocalization (Larson, 1991; Jürgens, 1994). The specific pattern of activation or suppression is determined by descending inputs from the ACG and limbic system, along with sensory feedback from the

![The Human Brain](image)
auditory, laryngeal and respiratory systems (Davis, Zhang, and Bandler, 1993; Ambalavanar et al., 1999). The resultant vocalizations convey the affective state of the organism. Although this system probably is responsible for emotional vocalizations in humans, it is unknown whether this pathway is involved in normal, nonemotive speech and language.

Neurons of the PAG project to several sites in the pons and medulla, one of which is the nucleus retroambiuous (NRA) (Holstege, 1989). The NRA in turn projects to the nucleus ambiguus (NA) and spinal cord motor neurons of the respiratory muscles. Lesions of the NRA eliminate vocalizations evoked by PAG stimulation (Shiba et al., 1997), and stimulation of the NRA elicits vocalization (Zhang, Bandler, and Davis, 1995). Thus, the NRA lies functionally between the PAG and motor neurons of laryngeal and respiratory muscles controlling vocalization and may play a role in coordinating these neuronal groups (Shiba et al., 1997; Luthe, Hausler, and Jürgens, 2000).

The PAG also projects to the parvocellular reticular formation, where neurons modulate their activity with temporal and acoustical variations in monkey calls, and lesions alter the acoustical structure of vocalizations (Luthe, Hausler, and Jürgens, 2000). These data suggest that the parvocellular reticular formation is important for the regulation of vocal quality and pitch.

Finally, the NA contains laryngeal motor neurons and is crucial to vocalization. Motor neurons in the NA control laryngeal muscles during vocalization, swallowing, and respiration (Yajima and Larson, 1993), and lesions of the NA abolish vocalizations elicited by PAG stimulation (Jürgens and Pratt, 1979). The NA receives projections either indirectly from the PAG, by way of the NRA (Holstege, 1989), or directly from the cerebral cortex (Kuypers, 1958). Sensory feedback for the reflexive control of laryngeal muscles flows through the superior and recurrent laryngeal nerves to the nucleus of the solitary tract and spinal nucleus of the trigeminal nerve (Tan and Lim, 1992).

In summary, vocalization is controlled by two pathways, one that is primitive and found in most animals, and one that is found only in humans and perhaps anthropoid apes (Fig. 2). The pathway found in all mammals extends from the ACG through the limbic system and midbrain PAG to medullary and spinal motor neurons, and seems to control most emotional vocalizations. Voluntary vocal control, found primarily in humans, aided by sensory feedback, may be exerted through a direct pathway from the motor cortex to the medulla. The tendency for vocalization and human speech to be strongly affected by emotions may suggest that all vocalizations rely at least in part on the ACG-PAG pathway. Details of how these two parallel pathways are integrated are unknown.

See also Vocal production system: Evolution.

—Charles R. Larson

References


Further Readings


**Voice Acoustics**

The basic acoustic source during normal phonation is a waveform consisting of a quasi-periodic sequence of pulses of volume velocity $U_s(t)$ that pass between the vibrating vocal folds (Fig. 1A). For modal vocal fold vibration, the volume velocity is zero in the time interval between the pulses, and there is a relatively abrupt discontinuity in slope at the time the volume velocity decreases to zero. The periodic nature of this waveform is reflected in the harmonic structure of the spectrum (Fig. 1B). The amplitudes of the harmonics at high frequencies decrease as $1/f^2$, where $f =$ frequency, i.e., at about $–12$ dB per octave. The frequency of this source waveform varies from one individual to another and within an utterance. In the time domain (Fig. 1A), a change in frequency is represented in the number of pulses per second; in the frequency domain (Fig. 1B), the frequency is represented by the spacing between the harmonics. The shape of the individual pulses can also vary with the speaker, and during an utterance the shape can be modified depending on the position within the utterance and the prominence of the syllable.

When the position and tension of the vocal folds are properly adjusted, a positive pressure below the glottis will cause the vocal folds to vibrate. As the cross-sectional area of the glottis changes during a cycle of vibration, the airflow is modulated. During the open phase of the cycle, the impedance of the glottal opening is usually large compared with the impedance looking into the vocal tract from the glottis. Thus, in most cases it is reasonable to represent the glottal source as a volume-velocity source that produces similar glottal pulses for different vocal tract configurations.

This source $U_s(t)$ is filtered by the vocal tract, as depicted in Figure 2. The volume velocity at the lips is $U_m(t)$, and the output sound pressure at a distance $r$ from the lips is $p_s(t)$. The magnitudes of the spectral components of $U_s(f)$ and $p_s(f)$ are shown below the corresponding waveforms in Figure 2. When a non-nasal vowel is produced, the vocal tract transfer function $T(f)$, defined as the ratio $U_m(f)/U_s(f)$, is an all-pole transfer function. The sound pressure $p_s$ is related to $U_m$ by a radiation characteristic $R(f)$. The magnitude of this radiation characteristic is approximately

$$|R(f)| = \frac{\rho \cdot 2\pi f}{4\pi r},$$

where $\rho =$ density of air. Thus we have

$$p_s(f) = U_s(f) \cdot T(f) \cdot R(f).$$

The magnitude of $p_s(f)$ can be written as

$$|p_s(f)| = |U_s(f) \cdot 2\pi f| \cdot |T(f)| \cdot \frac{\rho}{4\pi r}.$$  

The expression $|U_s(f) \cdot 2\pi f|$ is the magnitude of the Fourier transform of the derivative $U'_s(t)$. Thus the output sound pressure can be considered to be the result of filtering $U'_s(t)$ by the vocal tract transfer function $T(f)$, multiplied by a constant. That is, the derivative $U'_s(t)$ can be viewed as the effective excitation of the vocal tract.

For the ideal or modal volume-velocity waveform (Fig. 1), this derivative has the form shown in Figure 3A (Fant, Liljencrants, and Lin, 1985). Each pulse has a sequence of two components: (1) an initial smooth portion where the waveform is first positive, then passes through zero (corresponding to the peak of the pulse in Fig. 1), and then reaches a maximum negative value; and (2) a second portion where the waveform returns abruptly to zero, corresponding to the discontinuity in

![Figure 1](image1.png)

**Figure 1.** A, Idealized waveform of glottal volume velocity $U_s(t)$ for modal vocal fold vibration for an adult male speaker. B, Spectrum of waveform in A.
slope of the original waveform $U_s(t)$ at the time the vocal folds come together. The principal acoustic excitation of the vocal tract occurs at the time of this discontinuity. For this ideal or modal derivative waveform, the spectrum (Fig. 3B) at high frequencies decreases as $1/f$, i.e., at $-6 \text{ dB/octave}$, reflecting the discontinuity at closure.

For normal speech production, there are several ways in which the glottal waveform can differ from the modal waveform (or its derivative). One obvious attribute is the frequency $f_0$ of the glottal pulses, which is controlled primarily by changing the tension of the vocal folds, although the subglottal pressure also influences the frequency, particularly when the folds are relatively slack (Titze, 1989). Increasing or decreasing the subglottal pressure $P_s$ causes increases or decreases in the amplitude of the glottal pulses, or, more specifically, in the magnitude of the discontinuity in slope at the time of glottal closure. The magnitude of the glottal excitation increases roughly as $P_s^{3/2}$ (Ladefoged and McKinney, 1963; Isshiki, 1964; Tanaka and Gould, 1983).

Changes in the configuration of the membranous and cartilaginous portions of the vocal folds relative to the modal configuration can lead to changes in the waveform and spectrum of the glottal source. For some speakers and for some styles of speaking, the vocal folds and arytenoid cartilages are configured such that the glottis is never completely closed during a cycle of vibration, introducing several acoustic consequences. First, the speed with which the vocal folds approach the midline is reduced; the effect on the derivative waveform $U'_s(t)$ is that the maximum negative value is reduced (that is, it is less negative). Thus, the excitation of the vocal tract and the overall amplitude of the output are decreased. Second, there is continuing airflow throughout the cycle. The inertia of the air in the glottis and supraglottal airways prevents the occurrence of the

Figure 2. Schema showing how the acoustic source at the glottis is filtered by the vocal tract to yield a volume velocity $U_m(t)$ at the lips, which is radiated to obtain the sound pressure $p_r(t)$ at some distance from the lips. At the left of the figure both the source waveform $U_s(t)$ and its spectrum $U_s(f)$ are shown. At the right is the waveform $p_s(t)$ and spectrum $p_r(f)$ of the sound pressure. (Adapted with permission from Stevens, 1994.)

Figure 3. A, Derivative $U'_s(t)$ of the modal volume-velocity waveform in Figure 1. B, Spectrum of waveform in A.
abrupt discontinuity in $U_0'(t)$ that occurs at the time of vocal fold closure in modal phonation (Rothenberg, 1981). Rather, there is a non-zero return phase following the maximum negative peak, during which $U_0'(t)$ gradually returns to zero (Fant, Liljencrants, and Lin, 1985). The derivative waveform $U_0'(t)$ then has a shape that is schematized in Figure 4. The corresponding waveform $U_0(t)$ is shown below the waveform $U_0'(t)$. The spectral consequence of this non-zero return phase is a reduction in the high-frequency spectrum amplitude of $U_0(t)$ relative to the low-frequency spectrum amplitude. A third consequence of a somewhat abducted glottal configuration is an increased loss of acoustic energy from the vocal tract through the partially open glottis and into the subglottal airways. This energy loss affects the vocal tract filter rather than the source waveform. It is most apparent in the first formant range and results in an increased bandwidth of F1, causing a reduction in A1, the amplitude of the first-formant prominence in the spectrum (Hanson, 1997). The three consequences just described lead to a vowel for which the spectrum amplitude A1 in the F1 range is reduced and the amplitudes of the spectral prominences due to higher formants are reduced relative to A1.

Still another consequence of glottal vibration with a partially open glottis is that there is increased average airflow through the glottis, as shown in the $U_0(t)$ waveform in Figure 4. This increased flow causes an increased amplitude of noise generated by turbulence in the vicinity of the glottis. Thus, in addition to the quasi-periodic source, there is an aspiration noise source with a continuous spectrum (Klatt and Klatt, 1990). Since the flow is modulated by the periodic fluctuation in glottal area, the noise source is also modulated. This type of phonation has been called “breathy-voiced.”

The aspiration noise source can be represented as an equivalent acoustic volume-velocity source that is added to the periodic source. In contrast to the periodic source, the noise source has a spectrum that tilts upward with increasing frequency. It appears to have a broad peak at high frequencies, around 2–4 kHz (Stevens, 1998). Figure 5A shows estimated spectra of the periodic and noise components that would occur during modal phonation. The noise component is relatively weak, and is generated only during the open phase of glottal vibration. Phonation with a more abducted glottis of the type represented in Figure 4 leads to greater noise energy and reduced high-frequency amplitude of the periodic component, and the noise component may dominate the periodic component at high frequencies (Fig. 5B). With breathy-voiced phonation, the individual harmonics corresponding to the periodic component may be obscured by the noise component at high frequencies. At low frequencies, however, the harmonics are well defined, since the noise component is weak in this frequency region.

Figure 6 shows spectra of a vowel produced by a speaker with modal glottal vibration (A) and the same vowel produced by a speaker with a somewhat abducted

![Figure 4](image_url)  
*Figure 4. Schematized representation of volume velocity waveform $U_0(t)$ and its derivative $U_0'(t)$ when the glottis is never completely closed within a cycle of vibration.*

![Figure 5](image_url)  
*Figure 5. Schematized representation of spectra of the effective periodic and noise components of the glottal source for modal vibration (A) and breathy voicing (B). The spectrum of the periodic component is represented by the amplitudes of the harmonics. The spectrum of the noise is calculated with a bandwidth of about 300 Hz.*
Below the spectra are waveforms of the vowel before and after being filtered with a bandpass filter centered on F3, with a bandwidth of 600 Hz. The individual glottal pulses as filtered by F3 of the vowel are evident. B, Spectrum of the vowel /e/ produced by a male speaker who apparently phonated with a glottal chink. The waveforms below are as described in A. The noise in the waveform in the F3 region (and above) obscures the individual glottal pulses. The spectra are from Hanson and Chuang (1999). See text.

As the average glottal area increases, the transglottal pressure required to maintain vibration (phonation threshold pressure) increases. Therefore, for a given subglottal pressure an increase in the glottal area can lead to cessation of vocal fold vibration.

Figure 6. A, Spectrum of the vowel /e/ produced by a male speaker with approximately modal phonation. Below the spectrum are waveforms of this vowel before and after being filtered with a bandpass filter centered on F3, with a bandwidth of 600 Hz. The individual glottal pulses as filtered by F3 of the vowel are evident. B, Spectrum of the vowel /e/ produced by a male speaker who apparently phonated with a glottal chink. The waveforms below are as described in A. The noise in the waveform in the F3 region (and above) obscures the individual glottal pulses. The spectra are from Hanson and Chuang (1999). See text.
Figure 7. Distributions of H1*-A3*, a measure that reflects the reduction of the high-frequency spectrum relative to the low-frequency spectrum, for male (black bars) and female (gray bars) speakers. H1 is the amplitude of the first harmonic and A3 is the amplitude of the strongest harmonic in the F3 peak. The asterisks indicate that corrections have been applied to H1 and A3, as described in the text. (Adapted with permission from Hanson and Chuang, 1999.)


**Voice Disorders in Children**

Investigations of voice using aerodynamic techniques have been reported for more than 30 years. Investigators realized early on that voice production is an aero-mechanical event and that vocal tract aerodynamics reflect the interactions between laryngeal anatomy and complex physiological events. Aerodynamic events do not always have a one-to-one correspondence with vocal tract physiology in a dynamic biological system, but careful control of stimuli and a good knowledge of laryngeal physiology make airflow and air pressure measurements invaluable tools.

Airflow (rate of air movement or velocity) and air pressure (force per unit area of air molecules) in the vocal tract are good reflectors of vocal physiology. For example, at a simple level, airflow through the glottis (Vg) is an excellent indicator of whether the vocal folds are open or closed. When the vocal folds are open, there is airflow through the glottis, and when the vocal folds are completely closed, there is zero airflow. With other physiological events held constant, the amount of
airflow can be an excellent indicator of the degree of opening between the vocal folds.

Subglottal air pressure directly reflects changes in the size of the subglottal air cavity. A simplified version of Boyle’s law predicts the relationship in that a particular pressure \( P \) in a closed volume \( V \) of air must equal a constant \( K \), that is, \( K = PV \). Subglottal air pressure will increase when the size of the lungs is decreased; conversely, subglottal pressure will decrease when the size of the lungs is made larger. Changes in subglottal air pressure are mainly regulated through muscular forces controlling the size of the rib cage, with glottal resistance or glottal flow used to help increase or decrease the pressures (the glottis can be viewed as a valve that helps regulate pulmonary flows and pressures).

A small number of classic studies used average airflow and intraoral air pressure to investigate voice production in children (Subtelny, Worth, and Sakuda, 1966; Arkebauer, Hixon, and Hardy, 1967; Van Hattum and Worth, 1967; Beckett, Theolke and Cowan, 1971; Diggs, 1972; Bernthal and Beukelman, 1978; Stathopoulos and Weismer, 1986). Measures of flow and pressure were used to reflect laryngeal and respiratory function. During voice production, children produce lower average airflow than adults, and boys tend to produce higher average airflow than girls of the same age. Supraglottal and glottal airway opening most likely account for the different average airflow values as a function of age and sex. Assuming that pressure is the same across all speakers, a smaller supraglottal or glottal opening yields a higher resistance at the constriction and therefore a restricted or lower flow of air. The findings related to intraoral air pressure have indicated that children produce higher intraoral air pressures than adults, especially because they tend to speak at higher sound pressure levels (SPLs). The higher pressures produced by children versus adults reflect two physiological events. First, children tend to speak at a higher SPL than adults, and second, children’s airways are smaller and less compliant than adults’ (Stathopoulos and Weismer, 1986). Intuitively, it would appear that the greater peak intraoral air pressure in children should lead to a greater magnitude of oral airflow. It is likely that children’s smaller glottal and supraglottal areas substantially counteract the potentially large flows resulting from their high intraoral air pressures.

Children were found to be capable of maintaining the same linguistic contrasts as adults through manipulation of physiological events such as lung cavity size and driving pressure, and laryngeal and articulatory configuration. Other intraoral air pressure distinctions in children are similar to the overall trends described for adult pressures. Like adults, children produce higher pressures during (1) voiceless compared to voiced consonants, (2) prevocalic compared to postvocalic consonants, (3) stressed compared to unstressed syllables, and (4) stops compared to fricatives.

In the 1970s and 1980s, two important aerodynamic techniques relative to voice production were developed that stimulated new ways of analyzing children’s aero-

dynamic vocal function. The first technique was inverse filtering of the easily accessible oral airflow signal (Rothenberg, 1977). Rothenberg’s procedure allowed derivation of the glottal airflow waveform. The derived volume velocity waveform provides airflow values, permitting detailed, quantifiable analysis of vocal fold physiology. The measures made from the derived volume velocity waveform can be related to the speed of opening and closing of the vocal folds, the closed time of the vocal folds, the amplitude of vibration, the overall shape of the vibratory waveform, and the degree of glottal opening during the closed part of the cycle.

The second aerodynamic technique developed was for the estimation of subglottal pressure and laryngeal airway resistance (Rlaw) through noninvasive procedures (Loqvist, Carlborg, and Kitzing, 1982; Smitheran and Hixon, 1981). Subglottal air pressure is of primary importance, because it is responsible for generating the pressure differential causing vocal fold vibration (the pressure that drives the vocal folds). Subglottal pressure is also important for controlling sound pressure level and for contributing to changes in fundamental frequency—all factors essential for normal voice production. The estimation of Rlaw offers a more general interpretation of laryngeal dynamics and can be used as a screening measure to quantify values outside normal ranges of vocal function.

Measures made using the Smitheran and Hixon (1981) technique include the following:

1. **Average oral airflow**: Measured during the open vowel /a/ at midpoint to obtain an estimate of laryngeal airflow.

2. **Intraoral air pressure**: Measured peak pressure during the voiceless [p] to obtain an estimate of subglottal pressure.

3. **Estimated laryngeal airway resistance**: Calculated by dividing the estimated subglottal pressure by estimated laryngeal airflow. This calculation is based on analogy with Ohm’s law, \( R = V/I \), where \( R \) = resistance, \( V \) = voltage, and \( I \) = current. In the speech system, \( R \) = laryngeal airway resistance, \( V \) = subglottal pressure (P), and \( I \) = laryngeal airflow (V). Thus, \( R = P/V \).

Measures made using the derived glottal airflow waveform important to vocal fold physiology include the following (Holmberg, Hillman, and Perkell, 1988):

1. **Airflow open quotient**: This measure is comparable to the original open quotient defined by Timcke, von Leden, and Moore (1958). The open time of the vocal folds (defined as the interval of time between the instant of opening and the instant of closing of the vocal cords) is divided by the period of the glottal cycle. Opening and closing instants on the airflow waveform are taken at a point equal to 20% of alternating airflow (OQ-20%).

2. **Speed quotient**: The speed quotient is determined as the time it takes for the vocal folds to open divided by the time it takes for the vocal folds to close. Opening
and closing instants on the waveform are taken at a point equal to 20% of alternating airflow. The measure reflects how fast the vocal folds are opening and closing and the asymmetry of the opening and closing phases.

3. Maximum flow declination rate: The measure is obtained during the closing portion of the vocal fold cycle and reflects the fastest rate of airflow shut-off. Differentiating the airflow waveform and then identifying the greatest negative peak on differentiated waveform locates the fastest declination. The flow measure corresponds to how fast the vocal folds are closing.

4. Alternating glottal airflow: This measure is calculated by taking the glottal airflow maximum minus minimum. This measure reflects the amplitude of vibration and can reflect the glottal area during vibratory cycle.

5. Minimum flow: This measure is calculated by subtracting minimum flow from zero. It is indicative of airflow leak due to glottal opening during the closed part of the cycle.

Additional measures important to vocal fold physiology include the following:

6. Fundamental frequency: This measure is obtained from the inverse-filtered waveform by means of a peak-picking program. It is the lowest vibrating frequency of the vocal folds and corresponds perceptually to pitch.

7. Sound pressure level: This measure is obtained at the midpoint of the vowel from a microphone signal and corresponds physically to vocal intensity and perceptually to loudness.

Voice production arises from a multidimensional system of anatomical, physiological, and neurological components and from the complex coordination of these biological systems. Many of the measures listed above have been used to derive vocal physiology. Stathopoulos and Sapienza (1997) empirically explored applying objective voice measures to children’s productions and discussed the data relative to developmental anatomical data (Stathopoulos, 2000). From these cross-sectional data as a function of children’s ages, a clearer picture of child vocal physiology has emerged. Because the anatomical structure in children is constantly growing and changing, children continually alter their movements to make their voices sound “normal.” Figures 1 through 7 show cross-sectional vocal aerodynamic data obtained in children ages 4–14 years. One of the striking features that emerge from the aerodynamic data is the change in function at 14 years of age for boys. After that age, boys and men functionally group together, while women and children seem to have more in common aerodynamically and physiologically. The data are discussed in relation to their physiological implications.

Estimated subglottal pressure: Children produce higher subglottal pressures than adults, and all speakers produce higher pressures when they produce higher SPLs (Fig. 1). Anatomical differences in the upper and lower airway will affect the aerodynamic output of the vocal tract. The increased airway resistance in children could substantially increase tracheal pressures (Muller and Brown, 1980).

Airflow open quotient (OQ-20%): Open quotient has traditionally been very closely correlated with SPL. In adults, it is widely believed that as SPL increases, the open quotient decreases. That is, the vocal folds remain closed for a longer proportion of the vibratory cycle as vocal intensity increases. As seen in Figures 2A and 2B, which show data from a wide age span and both sexes, only adults and older teenagers produce lower open quotients for higher SPLs. It is notable that the younger children and women produce higher OQ-20%, indicating that the vocal folds are open for a longer proportion of the cycle than in men and older boys, regardless of vocal intensity.

Maximum flow declination rate (MFDR): Children and adults regulate their airflow shut-off through a combination of laryngeal and respiratory strategies. Their MFDRs range from about 250 cc/s/s for comfortable levels of SPL to about 1200 cc/s/s for quite high SPLs. In children and adults, MFDR increases as SPL increases (Fig. 3). Increasing MFDR as SPL increases affects the acoustic waveform by emphasizing the high-frequency components of the acoustic source spectra (Titze, 1988).

Alternating glottal airflow: Fourteen-year-old boys and men produce higher alternating glottal airflows than younger children and women during vowel production for the high SPLs (Fig. 4). We can interpret the flow data to indicate that older boys and men produce higher alternating glottal airflows because of their larger laryngeal structures and greater glottal areas. Additionally, men and boys increase their amplitude of vibration during the high SPLs more than women and children do. Greater SPLs result in greater lateral excursion of the vibrating vocal folds; hence the higher alternating glottal airflows for adults. Younger children also increase their
amplitude of vibration when they increase their SPL, and we would assume an increase in the alternating flow values. The interpretation is somewhat complicated by the fact that younger children and women have a shorter vocal fold length and smaller area (Flanagan, 1958), thereby limiting airflow through the glottis.

**Fundamental frequency:** As expected, older boys and men produce lower fundamental frequencies than women and younger children. An interesting result predicted by Titze’s (1988) modeling data is that the 4- and 6-year-olds produce unusually high $f_0$ values when they increase their SPL to high levels (Fig. 5). Changes in fundamental frequency are more easily effected by increasing tracheal pressure when the vocal fold is characterized by a smaller effective vibrating mass, as in young children ages 4–6 years.

**Laryngeal airway resistance:** Children produce voice with higher $R_{law}$ than 14-year-olds and adults, and all speakers increase their $R_{law}$ when increasing their SPL (Fig. 6). Since $R_{law}$ is calculated by dividing subglottal...
pressure by laryngeal airflow, the high Rlaw for high SPL is largely due to higher values of subglottal pressure, since the average glottal airflow is the same across age groups. A basic assumption needs to be discussed here, and that is, that glottal airflow will increase when subglottal air pressure increases if laryngeal configuration/resistance is held constant. The fact that subglottal pressure increases for high SPLs but flow does not increase clearly indicates that Rlaw must be increasing. Physiologically, the shape and configuration of the laryngeal airway must be decreasing in size to maintain the constant airflow in the setting of increasing subglottal pressures. In sum, children and adults alike continually modify their glottal airway to control the important variables of subglottal pressure and SPL.

The cross-sectional aerodynamic data, and in particular the flow data, make a compelling argument that the primary factor affecting children’s vocal physiology is the size of their laryngeal structure. A general scan of the cross-sectional data discussed here shows a change in vocal function at age 14 in boys. It is not merely coincidental that at 14 years, male larynges continue to increase in size to approximate the size of adult male larynges, whereas larynges in teenage girls plateau and approximate the size of adult female larynges (Fig. 7). Regardless of whether it is size or other anatomical factors affecting vocal function, it is clear that use of an adult male model for depicting normal vocal function is inappropriate for children. Age- and sex-appropriate aerodynamic, acoustic, and physiological models of normal voice need to be referred to for the diagnosis and remediation of voice disorders.

See also INSTRUMENTAL ASSESSMENT OF CHILDREN’S VOICE.

—Elaine T. Stathopoulos

References


**Voice Disorders of Aging**

Voice disorders afflict up to 12% of the elderly population (Shindo and Hanson, 1990). Voice disorders in elderly persons can result from normal age-related changes in the voice production mechanism or from pathological conditions separate from normal aging (Linville, 2001). However, distinguishing between pathology and normal age-related changes can be difficult. Indeed, a number of investigators have concluded that the vast majority of elderly patients with voice disorders suffer from a disease process associated with aging rather than from a disorder involving physiological aging alone (Morrison and Gore-Hickman, 1986; Woo et al., 1992). Carcinoma of the head and neck occasionally occurs late in life, although more commonly it is diagnosed between the ages of 50 and 70 (Leon et al., 1998). Interestingly, multiple etiologic factors related to a voice disorder are more common in elderly patients than in younger adults. In addition, elderly persons are at increased risk for laryngeal side effects from pharmaceutical agents, since prescription and nonprescription drugs are used disproportionately by the elderly (Linville, 2001).

Elderly patients often exhibit neurological voice disorders, particularly in later stages of old age. Estimates of the incidence of peripheral laryngeal nerve damage in elderly dysphonic patients range from 7% to 21%. Generally, peripheral paralysis in the elderly tends to be associated with disease processes associated with aging (such as lung neoplasm), as opposed to idiopathic peripheral paralysis, which occurs infrequently (Morrison and Gore-Hickman, 1986; Woo et al., 1992). Symptoms of peripheral paralysis include glottic insufficiency, reduced loudness, breathiness, and diplopaphonia. Voice therapy for peripheral paralysis frequently involves increasing vocal fold adductory force to facilitate closure of the glottis and improving breath support to minimize fatigue and improve speech phrasing. After age 60, central neurological disorders such as stroke, focal dystonia, Parkinson’s disease, Alzheimer’s disease, and essential tremor also occur frequently. Treatment for central disorders involves attention to specific deficits in vocal fold function such as positioning deficits, instability of vibration, and incoordination of movements. Functioning of the velopharynx, tongue, jaw, lips, diaphragm, abdomen, and rib cage may also be compromised and may require treatment. Treatment may focus on vocal fold movement patterns, postural changes, coordination of respiratory and phonatory systems, respiratory support, speech prosody, velopharyngeal closure, or speech intelligibility. In some cases, augmentative communication strategies might be used, or medical treatment may be combined with speech or voice therapy to improve outcomes (Ragig and Scherer, 1992).

A variety of benign vocal lesions are particularly prevalent in the elderly, including Reineke’s edema, polypoid degeneration, unilateral sessile polyp, and benign epithelial lesions with variable dysplastic changes (Morrison and Gore-Hickman, 1986; Woo et al., 1992). Reineke’s edema and polypoid degeneration occur more commonly in women and are characterized by chronic, diffuse edema extending along the entire length of the vocal fold. The specific site of the edema is the superficial layer of the lamina propria. Although the etiology of Reineke’s edema and polypoid degeneration is uncertain, reflux, cigarette smoking, and vocal abuse/misuse have been mentioned as possible causal factors (Koufman, 1995; Zeitels et al., 1997). Some degree of edema and epithelial thickening is a normal accompaniment of aging in some individuals. The reason why women are at greater risk than men for developing pathological epi-
therial changes as they age is unknown, although differences in vocal use patterns could be a factor. That is, elderly women may be more likely to develop hypertensive phonatory patterns in an effort to compensate for the age-related pitch lowering that accompanies vocal fold thickening and edema (Linville, 2001).

The incidence of functional hypertensive dysphonia among elderly speakers is disputed. Some investigators report significant evidence of phonatory behaviors consistent with hypertension, such as hyperactivity of the ventricular vocal folds in the elderly population (Hagen, Lyons, and Nuss, 1996; Morrison and Gore-Hickman, 1986). Others report a low incidence of vocal fold lesions commonly associated with hyperfunction (vocal nodules, pedunculated polyps), as well as relatively few cases of functional dysphonia without tissue changes (Woo et al., 1992). Clinicians are in agreement, however, that elderly patients need to be evaluated for evidence of hypertensive phonation and provided with therapy to promote more relaxed phonatory adjustments when evidence of hypertension is found, such as visible tension in the cervical muscles, a report of increased phonatory effort, a pattern of glottal attack, high laryngeal position, and/or anteroposterior laryngeal compression.

Inflammatory conditions such as pachydermia, laryngitis sicca, and nonspecific laryngitis also are diagnosed with some regularity in the elderly (see Infectious Diseases and Inflammatory Conditions of the Larynx). These conditions might arise as a consequence of smoking, reflux, medications, or poor hydration and often coexist with vocal fold lesions that may be either benign or malignant. Age-related laryngeal changes such as mucous gland degeneration might be a factor in development of laryngitis sicca (Morrison and Gore-Hickman, 1986; Woo et al., 1992). Gastroesophageal reflux disease (GERD) is another inflammatory condition that is reported to occur with greater frequency in the elderly (Richter, 2000). Since GERD has been present for a longer time in the elderly in comparison with younger adults, it is a more complicated disease in this group. Often elderly patients report less severe heartburn but have more severe erosive damage to the esophagus (Katz, 1998; Richter, 2000).

Because of advanced age, elderly patients may be at increased risk for traumatic injury to the vocal folds. Trauma might manifest as granuloma or scar tissue from previous surgical procedures requiring general anesthesia, or from other traumatic vocal fold injuries. Vocal fold scarring may be present as a consequence of previous vocal fold surgery, burns, intubation, inflammatory processes, or radiation therapy for glottic carcinoma (Morrison and Gore-Hickman, 1986; Kahane and Beckford, 1991).

Age-related changes in the endocrine system also affect the voice. Secretion disorders of the thyroid (both hyperthyroidism and hypothyroidism) occur commonly in the elderly and often produce voice symptoms, either as a consequence of altered hormone levels or as a result of increased pressure on the recurrent laryngeal nerve. In addition, voice changes are possible with thyroidec- tomy, even if the procedure is uncomplicated (e.g., Debruyne et al., 1997; Francis and Wartofsky, 1992; Sataloff, Emerich, and Hoover, 1997). Elderly persons also may experience vocal symptoms as a consequence of hypoparathyroidism or hyperparathyroidism, or from neuropathic disturbances resulting from diabetes (Maceri, 1986).

Lifestyle factors and variability among elderly speakers often blur the distinction between normal and disordered voice. Elderly persons differ in the rate and extent to which they exhibit normal age-related anatomical, physiological, and neurological changes. They also differ in lifestyle. These factors result in considerable variation in phonatory characteristics, articulatory precision habits, and respiratory function capabilities among elderly speakers (Linville, 2001).

Lifestyle factors can either postpone or exacerbate the effects of aging on the voice. Although a potentially limitless combination of environmental factors combine to affect aging, perhaps the most controllable and potentially significant lifestyle factors are physical fitness and cigarette smoking. The elderly population is extremely variable in fitness levels. The rate and extent of decline in motor and sensory performance with aging varies both within and across elderly individuals (Finch and Schneider, 1985). Declines in motor performance are directly related to muscle use and can be minimized by a lifestyle that includes exercise. The benefits of daily exercise include facilitated muscle contraction, enhanced nerve conduction velocity, and increased blood flow (Spirduso, 1982; Finch and Schneider, 1985; De Vito et al., 1997). A healthy lifestyle that includes regular exercise may also positively influence laryngeal performance, although a direct link has yet to be established (Ringel and Chodzko-Zajko, 1987). However, there is evidence that variability on measures of phonatory function in elderly speakers is reduced by controlling for a speaker’s physiological condition (Ramig and Ringel, 1983). Physical conditioning programs that include aerobic exercise often are recommended for aging professional singers to improve respiratory and abdominal conditioning and to avoid tremolo, as well as to improve endurance, accuracy, and agility. Physical conditioning is also important in nonsingers to prevent dysphonia in later life (Sataloff, Spiegel, and Rosen, 1997; Linville, 2001).

The effects of smoking coexist with changes related to normal aging in elderly smokers. Smoking amplifies the impact of normal age-related changes in both the pulmonary and laryngeal systems. Elderly smokers demonstrate accelerated declines in pulmonary function, even if no pulmonary disease is detected (Hill and Fisher, 1993; Lee et al., 1999). Smoking also has a definite effect on the larynx and alters laryngeal function. Clinicians must consider smoking history in assessing an elderly speaker’s voice (Linville, 2001).

Clinicians also must be mindful of the overall health status of older patients presenting with voice disorders. Elderly dysphonic patients often are in poor general health and have a high incidence of systemic illness. Pulmonary disease and hypertensive cardiac disease have been cited as particularly prevalent in elderly voice
patients (Woo et al., 1992). If multiple health problems are present, elderly dysphonic patients may be less compliant in following therapeutic regimens, or treatment for voice problems may need to be postponed. In general, the diagnosis and treatment of voice disorders are more complicated if multiple medical conditions are present (Linville, 2001).

—Sue Ellen Linville

References


Further Readings


Voice Production: Physics and Physiology

When the vocal folds are near each other, a sufficient transglottal pressure will set them into oscillation. This oscillation produces cycles of airflow that create the acoustic signal known as phonation, the voicing sound source, the voice signal, or more generally, voice. This article discusses some of the mechanistic aspects of phonation.

The most general expression of forces in the larynx dealing with motion of the vocal folds during phonation is

\[ F(x) = mx'' + bx' + kx, \]  

where \( F(x) \) is the air pressure forces on the vocal fold tissues, \( m \) is the mass of the tissue in motion, \( b \) is a viscous coefficient, \( k \) is a spring constant coefficient, \( x \) is the position of the tissue from rest, \( x' \) is the velocity of the tissue, and \( x'' \) is the acceleration of the tissue. In multmass models of phonation (e.g., Ishizaka and Flanagan, 1972), this equation is used for each mass proposed. Each term on the right-hand side characterizes forces in the tissue, and the left-hand side represents the external forces. This equation emphasizes the understanding that mass, viscosity, and stiffness each play a role in the motion (normal or abnormal) of the vocal folds, that these are associated with the acceleration, velocity, and displacement of the tissue, respectively, and they are balanced by the external air pressure forces acting on the vocal folds.

Glottal adduction has three parts. (1) How close the vocal processes are to each other determines the posterior prephonatory closeness of the membranous vocal folds. (2) The space created by the intercartilaginous glottis determines the "constant" opening there through which some or all of the DC (baseline) air will flow. (3) The closeness of the membranous vocal folds partly determines whether vocal fold oscillation can take place. To permit oscillation, the vocal folds must be within the phonatory adductory range (not too far apart, and yet not too overly compressed; Scherer, 1995), and the transglottal pressure must be at or greater than the phonatory threshold pressure (Titze, 1992) for the prevailing conditions of the vocal fold tissues and adduction.

The fundamental frequency \( F_0 \), related to the pitch of the voice, will tend to rise if the tension of the tissue in motion increases, and will tend to fall if the length, mass, or density increases. The most general expression to date for pitch control has been offered by Titze (1994), viz.,

\[ F_0 = (0.5/L)\sqrt{\left(s_p/\rho\right)\left(1 + (d_{TA}/d) \ast (s_{am}/s_p) \ast a_{TA}\right)^{0.5}}, \]

where \( L \) is the vibrating length of the vocal folds, \( s_p \) is the passive tension of the tissue in motion, \( d_{TA}/d \) is the ratio of the depth of the thyroarytenoid (TA) muscle in vibration to the total depth in vibration (the other tissue in motion is the more medial mucosal tissue), \( s_{am} \) is the maximum active stress that the TA muscle can produce, \( a_{TA} \) is the activity level of the TA muscle, and \( \rho \) is the density of the tissue in motion. In the vocal folds are lengthened by rotation of the thyroid and cricoid cartilages through the contraction of the cricothyroid (CT) muscles, the passive stretching of the vocal folds increases their passive tension, and thus \( L \) and \( s_p \) tend to counter each other, with \( s_p \) being more dominant (\( F_0 \) generally rises with vocal fold elongation). Increasing subglottal pressure increases the lateral amplitude of motion of the vocal folds, thus increasing \( s_p \) (via greater passive stretch; Titze, 1994) and \( d_{TA} \), thereby increasing...
F0. Increasing the contraction of the TA muscle (σ_{TA}) would tend to stiffen the muscle and shorten the vocal fold length (L), both of which would raise F0 but at the same time decrease the passive tension (σ_{p}) and the depth of vibration (d_v), which would decrease F0. Typically, large changes in F0 are associated with increased contraction of both the CT and TA muscles (Hirano, Ohala, and Vennard, 1969; Titze, 1994). Thus, the primary control for F0 is through the coordinative contraction of the TA and CT muscles, and subglottal pressure. F0 control, including the differentiated contraction of the complex TA muscle, anterior pull by the hyoid bone (Honda, 1983), cricoid tilt via tracheal pull (Sundberg, Leanderson, and von Euler, 1989), and the associations with adduction and vocal quality all need much study.

The intensity of voiced sounds, related to the loudness of the voice, is a combination and coordination of respiratory, laryngeal, and vocal tract aspects. Intensity increases with an increase in subglottal pressure, which itself depends on both lung volume reduction (an increase in air pressure in the lungs) and adduction of the vocal folds (which offers resistance to the flow of air from the lungs). An increase in the subglottal pressure during phonation can affect the cyclic glottal flow waveform (Fig. 1) by increasing its flow peak, increasing the maximum flow declination rate (MFDR, the maximum rate that the flow shuts off as the glottis is closing), and the sharpness of the baseline corner when the flow is near zero (or near its minimum value in the cycle). Greater peak flow, MFDR, and corner sharpness respectively increase the intensity of F0, the intensity of the first formant region (at least), and the intensity of the higherpartials (Fant, Liljencrants, and Lin, 1985; Gauffin and Sundberg, 1989). Glottal adduction level greatly affects the source spectrum or quality of the voice, increasing the negative slope of the spectrum as one changes voice production from highly compressed voice (a relatively flat spectrum) to normal adduction to highly breathy voice (a relatively steep spectrum) (Scherer, 1995). The vocal tract filter function will augment the spectral intensity values of the glottal flow source in the region of the formants (resonances), and will decrease their intensity values in the valleys of the resonant structure (Titze, 1994). The radiation away from the lips will increase the spectrum slope (by about 6 dB per octave).

Maintenance of vocal fold oscillation during phonation depends on the tissue characteristics mentioned above, as well as the changing shape of the glottis and the changing intraglottal air pressures during each cycle. During glottal opening, the shape of the glottis corresponding to the vibrating vocal folds is convergent (wider in the lower glottis, narrower in the upper glottis), and the pressures on the walls of the glottis are positive due to this shape and to the (always) positive subglottal pressure (for normal egressive phonation) (Fig. 2). This positive pressure separates the folds during glottal opening. During glottal closing, the shape of the glottis is divergent (narrower in the lower glottis, wider in the upper glottis), and the pressures on the walls of the lower glottis are negative because of this shape (Fig. 2), and negative throughout the glottis when there also is rarefaction (negative pressure) of the supraglottal region. This alternation in glottal shape and intraglottal pressures, along with the alternation of the internal forces of the vocal folds, maintains the oscillation of the vocal folds. The exact glottal shape and intraglottal pressure changes, however, need to be established in the human larynx for the wide range of possible phonatory and vocal tract acoustic conditions.

![Figure 1. One cycle of glottal airflow.](image)

**Figure 1.** One cycle of glottal airflow. Uac is the varying portion of the waveform, and Ud is the offset or bias flow. The flow peak is the maximum flow in the cycle, MFDR is the maximum flow declination rate (derivative of the flow), typically located on the right-hand side of the flow pulse, and the corner curvature at the end of the flow pulse describes how sharp the corner “shut-off” is. The flow peak, MFDR, and corner sharpness are all important for the spectral aspects of the flow pulse (see text).

![Figure 2. Pressure profiles within the glottis.](image)

**Figure 2.** Pressure profiles within the glottis. The upper trace corresponds to the data for a glottis with a 10° convergence and the lower trace to data for a glottis with a 10° divergence, both having a minimal glottal diameter of 0.04 cm (using a Plexiglas model of the larynx; Scherer and Shinwari, 2000). The transglottal pressure was 10 cm H2O in this illustration. Glottal entrance is at the minimum diameter position for the convergent glottis, and the pressures on the walls of the lower glottis are positive due to this shape and to the (always) positive subglottal pressure (for normal egressive phonation) (Fig. 2). This positive pressure separates the folds during glottal opening. During glottal closing, the shape of the glottis is convergent (narrower in the lower glottis, wider in the upper glottis), and the pressures on the walls of the lower glottis are negative because of this shape (Fig. 2), and negative throughout the glottis when there also is rarefaction (negative pressure) of the supraglottal region. This alternation in glottal shape and intraglottal pressures, along with the alternation of the internal forces of the vocal folds, maintains the oscillation of the vocal folds. The exact glottal shape and intraglottal pressure changes, however, need to be established in the human larynx for the wide range of possible phonatory and vocal tract acoustic conditions.
When the two medial vocal fold surfaces are not mirror images of each other across the midline, the geometric asymmetry creates different pressures on the two sides (i.e., pressure asymmetries; Scherer et al., 2001, 2002) and therefore different driving forces on the two sides. Also, if there is tissue asymmetry, that is, if the two vocal folds themselves do not have equal values of tension (stiffness) and mass, one vocal fold may not vibrate like the other one, creating roughness, subharmonics, and cyclic groupings (Isshiki and Ishizaka, 1976; Gerratt et al., 1988; Wong et al., 1991; Titze, Baken, and Herzel, 1993; Steinecke and Herzel, 1995).

Figure 3 summarizes some basic aspects of phona-
tion. The upper left suggests muscle contraction effects of vocal fold length (via CT and TA action), adduction (via TA, lateral cricoarytenoid, posterior cricoarytenoid, and interarytenoid muscle contraction), tension (via CT, TA, and adduction), and glottal shape (via vocal fold length, adduction, and TA rounding effect). When lung volume reduction is then employed, glottal airflow and subglottal pressure are created, resulting in motion of the vocal folds (if the adduction and pressure are sufficient), glottal flow resistance (transglottal pressure divided by the airflow), and the fundamental frequency (and pitch) of the voice. With the vocal tract included, the glottal flow is affected by the resonances of the vocal tract (pressures acting at the glottis level) and the invariance of the air of the vocal tract (to skew the glottal flow wave-
form to the right; Rothenberg, 1983), and the output spectra (quality) and intensity (loudness) result from the combination of the glottal flow, resonance, and radiation from the lips.

Many basic issues of glottal aerodynamics, aero-
acoustics, and modeling remain unclear for both normal and abnormal phonation. The glottal flow (the volume velocity flow) is considered a primary sound source, and the presence of the false vocal folds may interfere with the glottal jet and create a secondary sound source (Zhang et al., 2001). The turbulence and vorticities of the glottal flow may also contribute sound sources (Zhang et al., 2001). The false vocal folds themselves may contribute significant control of the flow resistance through the larynx, from more resistance (decreasing the flow if the false folds are quite close) to less resistance (increasing the glottal flow when the false folds are in an intermediate position) (Agarwal and Scherer, in press). Computer modeling needs to be practical, as in two-
mass modeling (Ishizaka and Flanagan, 1972), but also closer to physiological reality, as in finite element mod-
eling (Alipour, Berry, and Titze, 2000; Alipour and Scherer, 2000). The most complete approach so far is to combine finite element modeling of the tissue with computational fluid dynamics of the flow (to solve the Navier-Stokes equations; Alipour and Titze, 1996). However, we still need models of phonation that are helpful in describing and predicting subtle aspects of laryngeal function necessary for differentiating vocal pathologies, phonation styles and types, and approaches for phonosurgery, as well as for providing rehabilitation and training feedback for clients.

See also voice acoustics.

—Ron Scherer

References


**Voice Quality, Perceptual Evaluation of**

Voice quality is the auditory perception of acoustic elements of phonation that characterize an individual speaker. Thus, it is an interaction between the acoustic speech signal and a listener’s perception of that signal. Voice quality has been of interest to scholars for as long as people have studied speech. The ancient Greeks associated certain kinds of voices with specific character traits; for example, a nasal voice indicated a spiteful and immoral character. Ancient writers on oratory emphasized voice quality as an essential component of polished speech and described methods for conveying a range of emotions appropriately, for cultivating power, brilliance, and sweetness, and for avoiding undesirable characteristics like roughness, brassiness, or shrillness (see Laver, 1981, for review).

Evaluation of vocal quality is an important part of the diagnosis and treatment of voice disorders. Patients usually seek clinical care because of their own perception of a voice quality deviation, and most often they judge the success of treatment for the voice problem by improvement in their voice quality. A clinician may also judge success by documenting changes in laryngeal anatomy or physiology, but in general, patients are more concerned with how their voices sound after treatment. Researchers from other disciplines are also interested in measuring vocal quality. For example, linguists are interested in how changes in voice quality can signal changes in meaning; psychologists are concerned with the perception of emotion and other personal information encoded in voice; engineers seek to develop algorithms for signal compression and transmission that preserve voice quality; and law enforcement officials need to assess the accuracy of speaker identifications.

Despite this long intellectual history and the substantial cross-disciplinary importance of voice quality, measurement of voice quality is problematic, both clinically and experimentally. Most techniques for assessing voice quality fall into one of two general categories: perceptual assessment protocols, or protocols employing an acoustic or physiologic measurement as an index of quality. In perceptual assessments, a listener (or listeners) rates a voice on a numerical scale or a set of scales representing the extent to which the voice is characterized by critical aspects of voice quality. For example, Fairbanks (1960) recommended that voices be assessed on 5-point scales for the qualities harshness, hoarseness, and breathiness. In the GRBAS protocol (Hirano, 1981), listeners evaluate voices on the scales Grade (or extent of pathology), Roughness, Breathiness, Asthenicity (weakness or lack of power in the voice), and Strain, with each scale ranging from 0 (normal) to 4 (severely disordered). A recent revision to this protocol (Dejonckere et al., 1998) has expanded it to GIRBAS by adding a scale for Instability. Many other similar protocols have been proposed. For example, the Wilson Voice Profile System (Wilson, 1977) includes 7-point scales for laryngeal tone, laryngeal tension, vocal abuse, loudness, pitch, vocal inflections, pitch breaks, diplophonia (perception of two pitches in the voice), resonance, nasal emission, rate, and overall vocal efficiency. A 13-scale protocol proposed by Hammarberg and Gauffin (1995) includes scales for assessing aphony (lack of voice), breathiness, tension, laxness, creakiness, roughness, grattings, pitch instability, voice breaks, diplophonia, falsetto, pitch, and loudness. Even more elaborate protocols have been proposed by Geller (1988; 17 parameters) and Laver (approximately
50 parameters; e.g., Greene and Mathieson, 1989). Methods like visual-analog scaling (making a mark on an undifferentiated line to indicate the amount of a quality present) or direct magnitude estimation (assigning any number—as opposed to one of a finite number of scale values—to indicate the amount of a quality present) have also been applied in efforts to quantify voice quality. Ratings may be made with reference to “anchor” stimuli that exemplify the different scale values, or with reference to a listener’s own internal standards for the different levels of a quality.

The usefulness of such protocols for perceptual assessment is limited by difficulties in establishing the correct and adequate set of scales needed to document the sound of a voice. Researchers have never agreed on a standardized set of scales for assessing voice quality, and some evidence suggests that differences between listeners in perceptual strategies are so large that standardization efforts are doomed to failure (Kreiman and Gerratt, 1996). In addition, listeners are apparently unable to agree in their ratings of voices. Evidence suggests that on average, more than 60% of the variance in ratings of voice quality is due to factors other than differences between voices in the quality being rated. For example, scale ratings may vary depending on variable listener attention, difficulty isolating single perceptual dimensions within a complex acoustic stimulus, and differences in listeners’ previous experience with a class of voices (Kreiman and Gerratt, 1998). Evidence suggests that traditional perceptual scaling methods are effectively matching tasks, where external stimuli (the voices) are compared to stored mental representations that serve as internal standards for the various rating scales. These idiosyncratic, internal standards appear to vary with listeners’ previous experience with voices (Verdonck de Leeuw, 1998) and with the context in which a judgment is made, and may vary substantially across listeners as well as within a given listener. In addition, severity of vocal deviation, difficulty isolating individual dimensions in complex perceptual contexts, and factors like lapses in attention can also influence perceptual measures of voice (de Krom, 1994). These factors (and possibly others) presumably all add uncontrolled variability to scalar ratings of vocal quality, and contribute to listener disagreement (see Gerratt and Kreiman, 2001, for review).

In response to these substantial difficulties, some researchers suggest substituting objective measures of physiologic function, airflow, or the acoustic signal for these flawed perceptual measures, for example, using a measure of acoustic frequency perturbation as a de facto measure of perceived roughness (see ACOUSTIC ASSESSMENT OF VOICE). This approach reflects the prevailing view that listeners are inherently unable to agree in their perception of such complex auditory stimuli. Theoretical and practical difficulties also beset this approach. Theoretically, we cannot know the perceptual importance of particular aspects of the acoustic signal without valid measures of that perceptual response, because voice quality is by definition the perceptual response to a particular acoustic stimulus. Thus, acoustic measures that purport to quantify vocal quality can only derive their validity as measures of voice quality from their causal association with auditory perception. Practically, consistent correlations have never been found between perceptual and instrumental measures of voice, suggesting that such instrumental measures are not stable indices of perceived quality. Finally, correlation does not imply causality: simply knowing the relationship of an acoustic variable to a perceptual one does not necessarily illuminate its contribution to perceived quality. Even if an acoustic variable were important to a listener’s judgment of vocal quality, the nature of that contribution would not be revealed by a correlation coefficient. Further, given the great variability in perceptual strategies and habits that individual listeners demonstrate in their use of traditional rating scales, the overall correlation between acoustic and perceptual variables, averaged across samples of listeners and voices, fails to provide useful insight into the perceptual process. (See Kreiman and Gerratt, 2000, for an extended review of these issues.)

Gerratt and Kreiman (2001) proposed an alternative solution to this dilemma. They measured vocal quality by asking listeners to copy natural voice samples with a speech synthesizer. In this method, listeners vary speech synthesis parameters to create an acceptable auditory match to a natural voice stimulus. When a listener chooses the best match to a test stimulus, the synthesis settings parametrically represent the listener’s perception of voice quality. Because listeners directly compare each synthetic token they create with the target natural voice, they need not refer to internal standards for particular voice qualities. Further, listeners can manipulate acoustic parameters and hear the result of their manipulations immediately. This process helps listeners focus attention on individual acoustic dimensions, reducing the perceptual complexity of the assessment task and the associated response variability. Preliminary evaluation of this method demonstrated near-perfect agreement among listeners in their assessments of voice quality, presumably because this analysis-synthesis method controls the major sources of variance in quality judgments while avoiding the use of dubiously valid scales for quality. These results indicate that listeners do in fact agree in their perceptual assessments of pathological voice quality, and that tools can be devised to measure perception reliably. However, how such protocols will function in clinical (rather than research) applications remains to be demonstrated. Much more research is certainly needed to determine a meaningful, parsimonious set of acoustic parameters that successfully characterizes all possible normal and pathological voice qualities. Such a set could obviate the need for voice quality labels, allowing researchers and clinicians to replace quality labels with acoustic parameters that are causally linked to auditory perception, and whose levels objectively, completely, and validly specify the voice quality of interest.

—Bruce Gerratt and Jody Kreiman
References


Voice Rehabilitation After Conservation Laryngectomy

Partial or conservation laryngectomy procedures are performed not only to surgically remove a malignant lesion from the larynx, but also to preserve some functional valving capacity of the laryngeal mechanism. Retention of adequate valvular function allows conservation of some degree of vocal function and safe swallowing. As such, the primary goal of conservation laryngectomy procedures is cancer control and oncologic safety, with a secondary goal of maintaining upper airway sphincteric function and phonatory capacity post-surgery. However, conservation laryngectomy will always necessitate tissue ablation, with disruption of the vibratory integrity of at least one vocal fold (Bailey, 1981). From the standpoint of voice production, any degree of laryngeal tissue ablation has direct and potentially highly negative implications for the functional capacity of the postoperative larynx. Changes in laryngeal structure result in aerodynamic, vibratory, and ultimately acoustic changes in the voice signal (Berke, Gerratt, and Hanson, 1983; Rizer, Schecter, and Coleman, 1984; Doyle, 1994).

Vocal characteristics following conservation laryngectomy are a consequence of anatomical influences and the resultant physiological function of the post-surgical laryngeal sphincter, as well as secondary physiological compensation. In some instances, this level of compensation may facilitate the communicative process, but in other instances such compensations may be detri-
mental to the speaker’s communicative effectiveness (Doyle, 1997). Perceptual observations following a variety of conservation laryngectomy procedures have been diverse, but data clearly indicate perceived changes in voice quality, the degree of air leakage through the reconstructed laryngeal sphincter, the appearance of compensatory hypervalving, and other features (Blau-grund et al., 1984; Leeper, Heeneman, and Reynolds, 1990; Hoasjoe et al., 1992; Doyle et al., 1995; Keith, Leeper, and Doyle, 1996). Two factors in particular, glottic insufficiency and the relative degree of compli-
ance and resistance to airflow offered by the recon-
structed valve, appear to play a significant role in compensation behaviors influencing auditory-perceptual assessments of voice quality (Doyle, 1997). Excessive closure of the laryngeal mechanism at either glottic or supraglottic (or both) levels might decrease air escape, but may also create abnormalities in voice quality due to active (volitional) hyperclosure (Leeper, Heeneman, and Reynolds, 1990; Doyle et al., 1995; Keith, Leeper, and Doyle, 1996; Doyle, 1997). Similarly, volitional, compensatory adjustments in respiratory volume in an effort to drive a noncompliant voicing source characterized by postsurgical increases in its resistance to airflow may negatively influence auditory-perceptual judgments of the voice by listeners. This may then call attention to the voice, with varied degrees of social pen-

Further Readings


Further Readings


ality. In this regard, the ultimate postsurgical effects of conservation laryngectomy on voice quality may result in unique limitations for men and women, and as such require clinical consideration.

The clinical evaluation of individuals who have undergone conservation laryngectomy initially focuses on identifying behaviors that hold the greatest potential to negatively alter voice quality. Excessive vocal effort and a harsh, strained voice quality are commonly observed (Doyle et al., 1995). Standard evaluation may include videofluoroscopy (via both rigid and flexible endoscopy) and acoustic, aerodynamic, and auditory-perceptual assessment. Until recently, only limited comprehensive data on vocal characteristics of those undergoing conservation laryngectomy have been available. Careful, systematic perceptual assessment has direct clinical implications in that information from such an assessment will lead to the definition of treatment goals and methods of monitoring potential progress. A comprehensive framework for the evaluation and treatment of voice alterations in those who have undergone conservation laryngectomy is available (Doyle, 1977).

Depending on the auditory-perceptual character of the voice, the clinician should be able to discern functional (physiological) changes to the sphincter that may have a direct influence on voice quality. Those auditory-perceptual features that most negatively affect overall voice quality should form the initial targets for therapeutic intervention. For example, the speaker’s attempt to increase vocal loudness may create a level of hyperclosure that is detrimental to judgments of voice quality. Although the rationale for such “abnormal” behavior is easily understood, the speaker must understand the relative levels of penalty it creates in a communicative context. Further treatment goals should focus on (1) enhancing residual vocal functions and capacities, and (2) efforts to reduce or eliminate compensatory behaviors that negatively alter the voice signal (Doyle, 1997). Thus, primary treatment targets will frequently address changes in voice quality and/or vocal effort. Increased effort may be compensatory in an attempt to alter pitch or loudness, or simply to initiate the generation of voice. Clinical assessment should determine whether voice change is due to under- or overcompensation for the disrupted sphincter. Therefore, strategies for voice therapy must address changes in anatomical and physiological function, the contributions of volitional compensation, and whether changes in voice quality may be the result of multiple factors.

Voice therapy strategies for those who have undergone conservation laryngectomy have evolved from strategies used in traditional voice therapy (e.g., Colton and Casper, 1990; Boone and McFarland, 1994). Doyle (1997) has suggested that therapy following conservation laryngectomy should focus on “(1) smooth and easy phonation; (2) a slow, productive transition to voice generation at the initiation of voice and speech production, (3) increasing the length of utterance in conjunction with consistently easy phonation, and (4) control of speech rate via phrasing.” The intent is to improve vocal efficiency and generate the best voice quality without excessive physical effort. Clinical goals that focus on “easy” voice production without excessive speech rate are appropriate targets. Common facilitation methods may involve the use of visual or auditory feedback, ear training, and respiration training (Boone, 1977; Boone and McFarland, 1994; Doyle, 1997).

Maladaptive compensations following conservation laryngectomy often tend to be hyperfunctional behaviors. However, a subgroup of individuals may present with weak and inaudible voices because of pain or discomfort in the early postsurgical period. Such compensations may remain when the discomfort has resolved, and may result in perceptible limitations in verbal communication. In such cases of hypofunctional behavior, voice therapy is usually directed toward facilitating increased approximation of the laryngeal valve by means of traditional voice therapy methods (Boone, 1977; Colton and Casper, 1990). A weak voice requires the clinician to orient therapy tasks toward systematically increasing glottal resistance. Although a “rough” or “effortful” voice may be judged as abnormal, it may be preferable for some speakers when compared to a breathy voice quality. This is of particular importance when evaluating goals and potential voice outcomes relative to the speaker’s sex.

The physical and psychological demands placed on the patient during initial attempts at voicing might increase levels of tension that ultimately may reduce the individual’s phonatory capability. Those individuals who exhibit increased fundamental frequency, excessively aperiodic voices, or intermittent voice stoppages may be experiencing problems that result from postoperative physiological overcompensation because they are struggling to produce voice. Many individuals who have undergone conservative laryngectomy may demonstrate considerable effort during attempts at postsurgical voice production, particularly early during treatment. Because active glottic hypofunction is infrequently noted in those who have undergone conservative laryngectomy, clinical tasks that focus on reducing overcompensation (i.e., hyperfunctional closure) are more commonly used.

See also LARYNGECTOMY.

—Philip C. Doyle

References


Breathing—the mechanical process of moving air in and out of the lungs—plays an important role in both speech and voice production; however, the emphasis placed on breathing exercises relative to voice disorders in the published literature is mixed. A review of voice therapy techniques by Casper and Murray (2000) did not suggest any breathing exercises for voice disorders. Some books on voice and voice disorders have no discussion of changing breathing behavior relative to voice disorders (Case, 1984; Colton and Casper, 1996), others do (e.g., Aronson, 1980; Boone and McFarlane, 2000; Cooper, 1973; Stemple, Glaze, and Gerdeleman, 1995). Although breathing exercises are advocated by some, little is known about the role played by breathing, either directly or indirectly, in disorders of the voice. Reed (1980) noted the lack of empirical evidence that breathing exercises were useful in ameliorating voice disorders.

At present, then, there is a paucity of data on the relationship of breathing to voice disorders. The data that do exist generally describe the breathing patterns that accompany voice disorders, but there are no data on what kind of breathing behavior might contribute to voice disorders. For example, sapienza, stathopoulos, and brown (1997) studied breathing kinematics during reading in ten women with vocal fold nodules. They found that the women used more air per syllable, more lung volume per phrase, and initiated breath groups at higher lung volumes than women without vocal nodules. However, as the authors point out, the breathing behavior observed in the women with nodules was most likely in response to inefficient valving at the larynx and did not cause the nodules.

Normal Breathing

When assessing and planning therapy for disorders of the voice, it is important to know what normal function
is. Hixon (1975) provides a useful parameterization of breathing for speech that includes volume, pressure, flow, and body configuration or shape. For conversational speech in the upright body position, the following apply. Lung volume is the amount of air available for speaking or vocalizing. The volumes used for speech are usually within the midvolume range of vital capacity (VC), beginning at 60% VC and ending at around 40% VC (Hixon, 1973; Hixon, Mead, and Goldman, 1976; Hoit et al., 1989). This volume range is efficient and economical, in that extra effort is not required to overcome recoil forces (Hixon, Mead, and Goldman, 1976). Most lung volume exchange for speech is brought about by rib cage displacement and not by displacement of the abdomen (Hixon, 1973). The rib cage is efficient in displacing lung volume because it covers a greater surface area of the lungs, it consists of muscle fiber types that are able to generate fast and accurate pressure changes, and it is well endowed with spindle organs for purposes of sensory feedback. Pressure (translaryngeal pressure) is related to the intensity of the voice (Bouhuys, Proctor, and Mead, 1966). Pressure for conversational speech is typically around 4–7 cm H2O, or 0.4–0.7 kPa (Ishihiki, 1964; Statthopoulos and Sapienza, 1993). Pressure is generated by both muscular and inherent recoil forces, and the interworking of these forces depends on the level of lung volume (Hixon, Mead, and Goldman, 1976). Reference to flow is in the macro sense and denotes shorter inspiratory durations relative to longer expiratory durations. This difference in timing reflects the speaker’s desire to maintain the flow of speech in his or her favor. In the upright body position, body configuration refers to the size of the abdomen relative to the size of rib cage. For speech production in the upright body position, the abdomen is smaller and the rib cage is larger relative to relaxation (Hixon, Goldman, and Mead, 1973).

The upright body configuration provides an economical and efficient mechanical advantage to the breathing apparatus. The abdomen not only produces lung volume change in the expiratory direction, it also optimizes the function of the diaphragm and rib cage. It does so in two ways. First, inward abdominal placement lifts the diaphragm and the rib cage (Goldman, 1974). This action positions the expiratory muscle fibers of the rib cage and muscle fibers of the diaphragm on a more favorable portion of their length-tension curve. This allows quicker and more forceful contractions for both inspiration and expiration while using less neural energy. Second, this inward abdominal position as it is maintained provides a platform against which the diaphragm and rib cage can push in order to produce the necessary pressures and flows for speech. If the abdomen did not offer resistance to the rib cage during the expiratory phase of speech breathing, it would be forced outward and would move in a paradoxical manner during expiration. Paradoxical motion results in reduced economy of movement, or wasted motion. Thus, the pressure generated by the rib cage would alter the shape of the breathing apparatus and would not assist in developing as rapid and as large an alveolar pressure change (Hixon and Weismer, 1995).

Effects of Posture

Although much is known about speech breathing, little of this information has found its way into the clinical literature and been applied to voice therapy. As a result, only one breathing technique to improve voice production is usually described: The client is placed supine and increased outward movement of the abdomen is observed as the client breathes at rest.

The changes that occur in speech breathing with a switch from the upright position are numerous and reflect the different effects of gravity (see Hoit, 1995, for a comprehensive tutorial). In supine speech breathing involves approximately 20% less of VC. The change of body configuration from the upright position to the supine position means that rib cage volume decreases and abdominal volume increases. This modification changes the mechanics of the breathing muscles and requires a different motor control strategy for speech production. For example, in the supine position there is little or no muscular activity of the abdomen during speaking, whereas in the upright body position the abdominal muscles are quite active (Hixon, Mead, and Goldman, 1976; Hoit et al., 1988). In light of the mechanical and neural control issues discussed earlier, it seems unwarranted to position an individual supine to teach “natural” breathing for speech and voice. With regard to breathing at rest, it should be noted that this task is not specific to speech. Kelso, Saltzman, and Tuller (1986) hypothesize that the control of speech is task-specific. Hixon (1982) showed kinematic data from a patient with Friedreich’s ataxia whose abdominal wall was assumed to be paralyzed because it showed no inward displacement and no movement during speech. However, when this patient laughed, his abdominal wall was displaced inward and displayed a great amount of movement. As Hoit (1995) points out, it seems unlikely that techniques for changing breathing behavior learned in a resting, supine position would generalize to an upright body position. It seems curious why this technique is advocated. It may be that this technique does not change breathing behavior but is effective in relaxing individuals with voice disorders. Relaxation techniques have been advocated to reduce systemic muscular tension in individuals with voice disorders (Boone and McFarlane, 2000; Greene and Mathieson, 1989), and breathing exercises are known to be beneficial in reducing heart rate and blood pressure (Grossman et al., 2001; Han, Stegen, Valek, Clement, and Woestijne, 1996).

Other Approaches

If learning breathing techniques in the supine position is not useful, what else might be done with the breathing apparatus to ameliorate voice disorders? Perhaps modifying lung volume would be useful. Hixon and Putnam (1983) described breathing behavior in a 30-year-old woman (a local television broadcaster) with a functional voice problem. Although she had a normal voice and no positive laryngeal signs, audible inspiratory turbulence was evident during her broadcasts. Using breathing kinematic measurement techniques, Hixon and Putnam
found that this person spoke in the lower range of her VC, between 45% and 10%. They believed the noisy inspirations were due to the turbulence created by increased resistance in the lower airways that occurs at low lung volumes. Therefore, it is possible the telecaster’s noisy inspirations could have been eliminated or reduced if she were to produce speech at higher (more normal) lung volumes. However, the authors did not report any attempts to modify lung volume. Of note, the woman said that when she spoke at lower lung volumes, her voice sounded more authoritative, and that her voice seemed to be much lighter when she spoke at higher lung volumes.

When a person inspires to higher lung volumes, the downward movement of the diaphragm pulls on the trachea, and this pulling is believed to generate passive adductory forces on the vocal folds (Zenker and Zenker, 1960). Solomon, Garlitz, and Milbrath (2000) found that in men, there was a tendency for laryngeal airway resistance to be reduced during syllable production at high lung volumes compared with low lung volumes. Milstein (1999), using video-endoscopic and breathing kinematic analysis, found that at high lung volumes, the laryngeal area appeared more dilated and the larynx was in a lower vertical position in the neck than during phonation at lower lung volumes. Plassman and Lansing (1990) showed that with training, individuals can produce consistently higher lung volumes during inspiration.

Even after the call by Reed (1980) more than 20 years ago for more empirical data on breathing exercises to treat voice disorders, little has been done. More research in this area is decisively needed. Research efforts should focus first on how and whether abnormal breathing behavior contributes to voice disorders. Then researchers should examine techniques that are found to be effective for changing this abnormal breathing behavior—if it exists. Efficacy research is of great importance because of the reluctance of third-party insurers to cover voice disorders.

— Peter Watson

References


Voice Therapy: Holistic Techniques

Whenever a voice disorder is present, a change in the normal functioning of the physiology responsible for voice production may be assumed. These physiological events are measurable and may be modified by voice therapy. Normal voice production depends on a relative balance among airflow, supplied by the respiratory system; laryngeal muscle strength, balance, coordination, and stamina; and coordination among these and the supraglottic resonators (pharynx, oral cavity, nasal cavity). Any disturbance in the relative physiological balance of these vocal subsystems may lead to or be perceived as a voice disorder. Disturbances may occur in respiratory volume, power, pressure, and flow, and may manifest in vocal fold tone, mass, stiffness, flexibility, and approximation. Finally, the coupling of the supraglottic resonators and the placement of the laryngeal tone may cause or be implicated in a voice disorder (Titze, 1994).

The overall causes of vocal disturbances may be mechanical, neurological, or psychological. Whatever the cause, one management approach is direct modification of the inappropriate physiological activity through direct exercise and manipulation. When all three subsystems of voice are addressed in one exercise, this is considered holistic voice therapy. Examples of holistic voice therapy include Vocal Function Exercises (Stemple, Glaze, and Klaben, 2000), Resonant Voice Therapy (Verdolini, 2000), the Accent Method of voice therapy (Kotby, 1995; Harris, 2000), and the Lee Silverman Voice Treatment (Ramig, 2000). The following discussion considers the use of Vocal Function Exercises to strengthen and balance the vocal mechanism.

The Vocal Function Exercise program is based on an assumption that has not been proved empirically. Nonetheless, this assumption and the clinical logic that follows have been supported through many years of clinical experience and observation, as well as several efficacy studies (Stemple, 1994; Sabol, Lee, and Stemple, 1995; Roy et al., 2001). In a double-blind,
placebo-controlled study, Stemple et al. (1994) demonstrated that Vocal Function Exercises were effective in enhancing voice production in young women without vocal pathology. The primary physiological effects were reflected in increased phonation volumes at all pitch levels, decreased airflow rates, and a subsequent increase in maximum phonation times. Frequency ranges were extended significantly in the downward direction.

Sabol, Lee, and Stemple (1995), experimenting with the value of Vocal Function Exercises in the practice regimen of singers, used graduate students of opera as subjects. Significant improvements in the physiologic measurements of voice production were achieved, including increased airflow volume, decreased airflow rates, and increased maximum phonation time, even in this group of superior voice users.

Roy et al. (2001) studied the efficacy of Vocal Function Exercises in a population with voice pathology. Teachers who reported experiencing voice disorders were randomly assigned to three groups: Vocal Function Exercises, vocal hygiene, and control groups. For 6 weeks the experimental groups followed their respective therapy programs and were monitored by speech-language pathologists trained by the experimenters in the two approaches. Pre- and post-testing of all three groups using the Voice Handicap Index (VHI; Jacobson et al., 1997) revealed significant improvement in the Vocal Function Exercise group, and no improvement in the vocal hygiene group. Subjects in the control group rated themselves worse.

The laryngeal mechanism is similar to other muscle systems and may become strained and imbalanced for a variety of reasons (Saxon and Schneider, 1995). Indeed, the analogy that we often draw with patients is a comparison of knee rehabilitation with rehabilitation of the voice. Both the knee and the larynx consist of muscle, cartilage, and connective tissue. When the knee is injured, rehabilitation includes a short period of immobilization to reduce the effects of the acute injury. The immobilization is followed by assisted ambulation, and then the primary rehabilitation begins, in the form of systematic exercise. This exercise is designed to strengthen and balance all of the supportive knee muscles for the purpose of returning the knee as close to its normal functioning as possible.

Rehabilitation of the voice may also involve a short period of voice rest after acute injury or surgery to permit healing of the mucosa to occur. The patient may then begin conservative voice use and follow through with all of the management approaches that seem necessary. Full voice use is then resumed quickly, and the therapy program often is successful in returning the patient to normal voice production. Often, however, patients are not fully rehabilitated because an important step was neglected—the systematic exercise program to regain the balance among airflow, laryngeal muscle activity, and supraglottic placement of the tone.

A series of laryngeal muscle exercises was first described by Bertram Briess (1957, 1959). Briess suggested that for the voice to be most effective, the intrinsic muscles of the larynx must be in equilibrium. Briess’s exercises concentrated on restoring the balance in the laryngeal musculature and decreasing tension of the hyperfunctioning muscles. Unfortunately, many assumptions Briess made regarding laryngeal muscle function were incorrect, and his therapy methods were not widely followed. The concept of direct exercise to strengthen voice production persisted. Barnes (1977) described a modification of Briess’s work that she termed Briess Exercises. These exercises were modified and expanded by Stemple (1984) into Vocal Function Exercises. The exercise program strives to balance and strengthen the subsystems of voice production, whether the disorder is one of vocal hyperfunction or hypofunction.

The exercises are simple to teach and, when presented appropriately, seem reasonable to patients. Indeed, many patients are enthusiastic to have a concrete program, similar in concept to physical therapy, during which they may plot the progress of their return to vocal efficiency. The program begins by describing the problem to the patient, using illustrations as needed or the patient’s own stroboscopic evaluation video. The patient is then taught a series of four exercises to be practiced at home, two times each, twice a day, preferably morning and evening. These exercises include the following:

1. Sustain the /i/ vowel for as long as possible on a musical note: F above middle C for females and boys, F below middle C for males. (Pitches may be modified up or down to fit the needs of the patient. Seldom are they modified by more than two scale steps in either direction.)

   The goal of the exercise is based on airflow volume. In our clinic, the goal is based on reaching 80–100 mL/s of airflow. So, if the flow volume is 4000 mL, the goal is 40–45 s. When airflow measurements are not available, the goal is equal to the longest /s/ that the patient is able to sustain. Placement of the tone should be in an extreme forward focus, almost but not quite nasal. All exercises are produced as softly as possible, but not breathy. The voice must be engaged. This is considered a warm-up exercise.

2. Glide from your lowest note to your highest note on the word knoll.

   The goal is to achieve no voice breaks. The glide requires the use of all laryngeal muscles. It stretches the vocal folds and encourages a systematic, slow engagement of the cricothyroid muscles. The word knoll encourages a forward placement of the tone as well as an expanded open pharynx. The patient’s lips are to be rounded, and a sympathetic vibration should be felt on the lips. (A lip trill, tongue trill, or the word whoop may also be used.) Voice breaks will typically occur in the transitions between low and high registers. When breaks occur, the patient is encouraged to continue the glide without hesitation. When the voice breaks at the top of the current range and the patient typically has more range, the glide may be continued without voice as the
folds will continue to stretch. Glides improve muscular control and flexibility. This is considered a stretching exercise.

3. Glide from your highest note to your lowest note on the word _knoll._

The goal is to achieve no voice breaks. The patient is instructed to feel a half-yawn in the throat throughout this exercise. By keeping the pharynx open and focusing the sympathetic vibration at the lips, the downward glide encourages a slow, systematic engagement of the thyroarytenoid muscles without the presence of a back-focused growl. In fact, no growl is permitted. (A lip trill, tongue trill, or the word _boom_ may also be used.) This is considered a contracting exercise.

4. Sustain the musical notes C-D-E-F-G for as long as possible on the word _knoll_ minus the _kn_. The range should be around middle C for females and boys, an octave below middle C for men.

The goal is the same as for exercise 1. The _-oll_ is produced with an open pharynx and constricted, sympathetically vibrating lips. The shape of the pharynx in respect to the lips is like an inverted megaphone. This exercise may be tailored to the patient’s present vocal ability. Although the basic range of middle C (an octave lower for men) is appropriate for most voices, the exercises may be customized up or down to fit the current vocal condition or a particular voice type. Seldom, however, is the exercise shifted more than two scale steps in either direction. This is considered a low-impact adductory power exercise.

The quality of the tone is also monitored for voice breaks, wavering, and breathiness. Tone quality improves as times increase and pathologic conditions begin to resolve. All exercises are done as softly as possible. It is much more difficult to produce soft tones; therefore, the vocal subsystems will receive a better workout than if louder tones are produced. Extreme care is taken to teach the production of a forward tone that lacks tension. In addition, attention is paid to the glottal onset of the tone. The patient is asked to breathe in deeply, with attention paid to training abdominal breathing, posturing the vowel momentarily, and then initiating the exercise gesture without a forceful glottal attack or an aspirated breathy attack. It is explained to the patient that maximum phonation times increase as the efficiency of the vocal fold vibration improves. Times do not increase with improved lung capacity. (Even aerobic exercise does not improve lung capacity, but rather the efficiency of oxygen exchange with the circulatory system, thus giving the sense of more air.)

The musical notes are matched to the notes produced by an inexpensive pitch pipe that the patient purchases for use at home, or a tape recording of live voice doing the exercises may be given to the patient for home use. Many patients find the tape-recorded voice easier to match than the pitch pipe. We have found that patients who think they are “tone deaf” can often be taught to approximate the correct notes well with practice and guidance from the voice pathologist.

Finally, patients are given a chart on which to mark their sustained times, which is a means of plotting progress. Progress is monitored over time and, because of normal daily variability, patients are encouraged not to compare today with tomorrow, and so on. Rather, weekly comparisons are encouraged. The estimated time of completion for the program is 6–8 weeks. Some patients experience minor laryngeal aching for the first day or two of the program, similar to the muscle aching that might occur with any new muscular exercise. As this discomfort will soon subside, they are encouraged to continue the program through the discomfort should it occur.

When the patient has reached the predetermined therapy goal, and the voice quality and other vocal symptoms have improved, then a tapering maintenance program is recommended. Although some professional voice users choose to remain in peak vocal condition, many of our patients desire to taper the exercise program. The following systematic taper is recommended:

- Full program, 2 times each, 2 times per day
- Full program, 2 times each, 1 time per day (morning)
- Full program, 1 time each, 1 time per day (morning)
- Exercise 4, 2 times each, 1 time per day (morning)
- Exercise 4, 1 time each, 1 time per day (morning)
- Exercise 4, 1 time each, 3 times per week (morning)
- Exercise 4, 1 time each, 1 time per week (morning)

Each taper should last 1 week. Patients should maintain 85% of their peak time; otherwise they should move up one step in the taper until the 85% criterion is met.

Vocal Function Exercises provide a holistic voice treatment program that attends to the three major subsystems of voice production. The program appears to benefit patients with a wide range of voice disorders because it is reasonable in regard to time and effort. It is similar to other recognizable exercise programs: the concept of physical therapy for the vocal folds is understandable; progress may be easily plotted, which is inherently motivating; and it appears to balance and strengthen the relationships among airflow, laryngeal muscle activity, and supraglottic placement.

—Joseph Stempel

References

Barnes, J. (1977, October). Briess exercises. Workshop presented at the Southwestern Ohio Speech and Hearing Association, Cincinnati, OH.


Voice Therapy for Adults

Voice therapy for adults may be motivated by functional, health-related, or diagnostic considerations. Functional issues are the usual indication. Adults with voice problems often experience significant functional disruptions in occupational, social, communicative, physical, or emotional domains, and in selected populations, voice therapy is effective in reducing such disruptions. Health-related concerns are less common precipitants of voice therapy in adults. However, physical disease such as cancerous, precancerous, inflammatory, or neurogenic disease may exist and may be exacerbated by behavioral factors such as smoking, diet, hydration, or phonotrauma. Voice therapy may be a useful adjunct to medical or surgical treatment in these cases. Finally, voice therapy may be indicated in cases of diagnostic uncertainty. A classic situation is the need to distinguish between functional and neurogenic conditions. The restoration of a normal or near-normal voice with therapy may suggest a functional origin of the problem. Lack of voice restoration suggests the need for further clinical studies to rule out neurological causes.

Voice therapy can be characterized with reference to several different classification schemes, which results in a certain amount of nosological confusion. Many of the conditions listed in the various classifications map to several different voice therapy options, and by the same token, each therapy option maps to multiple classifications. Here we review voice therapy in relation to (1) vocal biomechanics and (2) a specific therapy approach—roughly the “what” and “how” of voice therapy.

**Vocal Biomechanics.** The preponderance of voice problems that are amenable to voice therapy involve some form of abnormality in vocal fold adduction. Phonotraumatic lesions such as nodules, polyps, and non-specific inflammation consequent on voice use are traceable to hyperadduction resulting from vocal fold impact stress. Adduction causes monotonic increases in impact stress. Adduction causes monotonic increases in impact stress (Jiang and Titze, 1994). In turn, impact stress appears to be a primary cause of phonotrauma (Titze, 1994). Thus, therapy targeting a reduction in adduction is indicated in cases of hyperadduction. Another large group of diagnostic conditions involves hypoadduction of the vocal folds. Examples include vocal fold paralysis, paresis, atrophy, bowing, and non-added hyperfunction (muscle tension dysphonia; for a discussion, see Hillman et al., 1989). Treatment that increases vocal fold closure is indicated in such cases.

Voice therapy addresses adductory deviations using a variety of biomechanical solutions. The traditional approach to hyperadduction and its sequelae has targeted the use of widely separated vocal folds and small-amplitude oscillations during voice production; examples are use of a “quiet, breathy voice” (Casper et al., 1989; Casper, 1993) or quiet “yawn-sigh” phonation (Boone and McFarlane, 1993). This general approach is sensible for the reduction of hyperadduction and thus phonotraumatic changes, in that vocal fold impact stress, and phonotrauma, should be reduced by it. There is evidence that the quiet, breathy voice approach is effective in reducing signs and symptoms of phonotraumatic lesions for individuals who use it outside the clinic (Verdolini-Marston et al., 1995). However, individuals may also restrict their use of a quiet, breathy voice extraclinically because it is functionally limiting (Verdolini-Marston et al., 1995).
The traditional approach to hypoadduction has involved “pushing” and “pulling” exercises, which should reduce the glottal gap (e.g., Boone and McFarlane, 1994). Indeed, some data corroborate clinicians’ impressions that this approach can increase voice intensity in individuals with glottal incompetence (Yamaguchi et al., 1993).

A more recent approach to treating adductory abnormalities has focused on the use of a single “ideal” vocal fold configuration as the target for both hyperadduction and hypoadduction. The configuration involves barely separated vocal folds, which is “ideal” because it optimizes the trade-off between voice output strength (relatively strong) and vocal fold impact stress, and thus reduces the potential for phonotraumatic injury (Berry et al., 2001). Voice produced with this intermediate laryngeal configuration has been called “resonant voice,” perceptually corresponding to anterior oral vibratory sensations during “easy” voicing (Verdolini et al., 1998). Programmatic approaches to resonant voice training have shown reductions in phonatory effort, voice quality, and laryngeal appearance (Verdolini-Marston et al., 1995), as well as reductions in functional disruptions due to voice problems in individuals with conditions known or presumed to be related to hyperadduction, such as nodules. Moreover, there is evidence that individuals use this type of voicing outside the clinic more than the traditional “quiet, breathy voice” because it is functionally tractable (Verdolini-Marston et al., 1995). Resonant voice training may also be useful in improving vocal and functional status in individuals with hypoadducted dysphonia. Recent theoretical modeling has indicated that nonlinear source (vocal fold)–filter (vocal tract) interactions are critical in maximizing voice output germane to resonant voice and other voice types (Titze, 2002).

A relatively small number of clinical cases involve vocal fold elongation abnormalities as the salient feature of the vocal condition. Often, the medical condition involves cricothyroid paresis, although thyroarytenoid paresis may also be implicated. Voice therapy has been less successful in treating such conditions. Other elongation abnormalities are functional, as in mutational falsetto. The clinical consensus is that voice therapy generally is useful in treating mutational falsetto.

Finally, in addition to addressing laryngeal kinematics, voice therapy usually also addresses nonphonatory aspects of biomechanics that influence the vocal fold mucosa. Such issues are addressed in voice hygiene programs (see VOICE HYGIENE). Mucosal performance and mucosal vulnerability to trauma are the key concerns. The primary issues targeted are hydration and behavioral control of laryngopharyngeal reflex. Dehydration increases the pulmonary effort required for phonation, whereas hydration decreases it and also decreases laryngeal phonotrauma (e.g., Titze, 1988; Verdolini, Titze, and Fennell, 1994; Solomon and DiMattia, 2000). Thus, hydration regimens are appropriate for individuals with voice problems and dehydration (Verdolini-Marston, Sandage, and Titze, 1994). There is increasing support for the view that laryngopharyngeal reflex plays a role in a wide range of laryngeal diseases, including inflammatory and even neurogenic and malignant disease. Voice therapy can play a supportive role to the medical or surgical treatment of laryngopharyngeal reflux by educating patients regarding behavioral issues such as diet and sleeping position. Some data are consistent with the view that control of laryngopharyngeal reflux can improve both laryngeal appearance and voice symptoms in individuals with a diagnosis of laryngopharyngeal reflux (Shaw et al., 1996; Hamdan et al., 2001). However, vocal hygiene programs alone in voice therapy apparently produce little benefit if they are not coupled with voice production work.

**Specific Therapy Approach.** Recently, interest has emerged in cognitive mechanisms involved in skill acquisition and factors affecting patient compliance as related to voice training and therapy models. Speech-language pathologists may train individuals to acquire the basic biomechanical changes described in preceding paragraphs, and others. The traditional approach is eclectic and entails implementing a series of facilitating techniques such as the “yawn-sigh” and “push-pull” techniques, as well as other maneuvers, such as altering the tongue position, changing the loudness of the voice, using chant talk, and using digital manipulation. Facilitating techniques are used by many clinicians and are generally considered effective. However, formal efficacy data are lacking for most of the techniques. An exception is digital manipulation, specifically manual circumlaryngeal therapy (laryngeal massage), used for idiopathic, presumably hyperfunctional dysphonia. Brief courses of aggressive laryngeal massage by skilled practitioners have dramatically improved voice in individuals with this condition (Roy et al., 1997). Also, variants of “yawn-sigh” phonation, such as falsetto and breathy voicing, may temporarily improve symptoms of adductor spasmodic dysphonia and increase the duration of the effectiveness of botulinum toxin injections (Murry and Woodson, 1995).

Several programmatic approaches to voice therapy have been developed, some of which have been submitted to formal clinical studies. An example is the Lee Silverman Voice Treatment (LSVT). This treatment uses “loud” voice to treat not only hypoadduction and hypophonia, but also prosodic and articulatory deficiencies in individuals with Parkinson’s disease. LSVT utilizes a predetermined hierarchy of speech tasks in 16 therapy sessions delivered over 4 weeks. In comparison with control and alternative treatment groups, LSVT has increased vocal loudness and voice inflection for as long as 2 years following therapy termination (Ramig, Sapir, Fox et al., 2001; Ramig, Sapir, Countrymen et al., 2001). Critical aspects of LSVT that may contribute to its success include a large number of repetitions of the target “loud voice” in a variety of physical contexts.

Another programmatic approach to voice therapy, the Lessac-Madsen Resonant Voice Therapy (LMRVT), was developed for individuals with either hyper- or hypoadducted voice problems associated with nodules,
polyps, nonspecific phonotraumatic changes, paralysis, paresis, atrophy, bowing, and sulcus vocalis. LMRVT targets the use of barely touching or barely separated vocal folds for phonation, a configuration considered to be ideal because it maximizes the ratio of voice output intensity to vocal fold impact intensity (Berry et al., 2001). In LMRVT, eight structured therapy sessions typically are delivered over 8 weeks. Training emphasizes sensory processing and the extension of “resonant voice” to a variety of communicative and emotional environments. Data on preliminary versions of LMRVT indicate that it is as useful as quiet, breathy voice training for sorority women with phonotrauma or the use of amplification for teachers with voice problems in reducing various combinations of phonatory effort, voice quality, laryngeal appearance, and functional status (Verdolini-Marston et al., 1995).

Another programmatic approach to voice therapy for both hyper- and hypoadducted conditions is called Vocal Function Exercises (VFE; Stemple et al., 1994). This approach targets similar vocal fold biomechanics as LMRVT, that is, vocal folds that are barely touching or barely separated, for phonation. Training consists of repeating maximally sustained vowels and pitch glides twice daily over a period of 4–6 weeks. Carryover exercises to conversational speech may also be used. A 6-week program of VFE in teachers with voice problems resulted in greater self-perceived voice improvement, greater phonatory ease, and better voice clarity than that achieved with vocal hygiene treatment alone (Roy et al., 2001).

Another program, Accent Therapy, addresses the ideal laryngeal configuration—barely touching or barely separated vocal folds—in individuals with hyper- and hypoadducted conditions (Smith and Thyme, 1976). Training entails the use of specified rhythmic, prosodically stressed vocal repetitions, beginning with sustained consonants and progressing to phrases and extended speech. The Accent Method is more widely used in Europe and Asia than in the United States.

Electromyographic biofeedback has been reported to be effective in reducing laryngeal hyperfunction and laryngeal appearance in individuals with voice problems linked to hyperadduction (nodules). Also, visual feedback using videendoscopy may be useful in treating numerous voice conditions; specific clinical observations have been reported relative to ventricular phonation (Bastian, 1987).

Finally, some clinicians have found that sensory differentiation exercises may help in the treatment of repetitive strain injury—one of the fastest growing occupational injuries. Repetitive strain injury involves decreased use of manual digits or voice and pain subsequent to overuse. Attention to sensory differentiation in the treatment of repetitive strain injury is motivated by reports of fused representation for groups of movements in sensory cortex following extensive digit use (e.g., Byl, Merzenich, and Jenkins, 1996).

—Katherine Verdolini

References


The neurological impact to the aging larynx includes central and peripheral motor nervous system changes. Central nervous system changes include nerve cell losses in the cortex of the frontal, parietal, and temporal lobes of the brain. This results in the slowing of motor movements (Scheibel and Scheibel, 1975). Nerve conduction velocity also contributes to speed of voluntary movements such as pitch changes, increased loudness, and speed of articulation (Leonard et al., 1997). Nervous system changes are also associated with tremor, a
condition seen more in the elderly than in young individuals. Finally, dopaminergic changes which decline with aging may also affect the speed of motor processing (Morgan and Finch, 1988).

The peripheral changes that occur in the elderly are thought to be broadly related to environmental effects of trauma (Woo et al., 1992), selective denervation of type II fast twitch muscle fibers (Lexell, 1997) and decrease in distal and motor neurons, resulting in decreased contractile strength and an increase in muscle fatigue (Doherty and Brown, 1993).

**Voice Changes Related to Neurological Aging**

The central and peripheral degeneration and concomitant regenerative neural changes that occur with neurological aging may result in a number of voice disorders. Excluding vocal fold paralysis, these neurological changes account for disorders of voice quality and overall vocal output. Murry and Rosen reported on 205 patients 65 years of age and older. Table 1 shows the diagnosis of this group (Murry and Rosen, 1999).

The most common symptoms reported by this group of patients are shown in Table 2.

Neurological changes to the voice accompanying aging are related to decreased neurological structure and function which result in patient perceived and listener perceived vocal dysfunction. Indeed, if the neuromotor systems are intact and the elderly patient is healthy, the speaking and singing voice is not likely to be perceived as "old" nor function as "old" (McGlone and Hollien, 1963). Conversely, the voice may be perceived as "old" not solely due to neurological changes in the larynx and upper airway, but due to muscular weakness of the upper body (Ramig and Ringel, 1983), cardiovascular changes (Orlikoff, 1990), or decreased hearing acuity resulting in excessive vocal force (Chodzko-Zajko and Ringel, 1987).

There are, however, certain aspects of voice production that are characteristically associated with age-related neuropathy. The clinical examination of elderly individuals who complain of voice disorders should specifically address and test for loss of vocal range and volume, vocal fatigue, increased breathy quality during extended conversations, presence of tremor, and pitch breaks (especially breaks into falsetto and hoarse voice quality). Elderly singers should be evaluated for pitch inaccuracies, increased breathiness, and changes in vibrato (Tanaka, Hiro, and Chijina, 1994).

A careful examination of the elderly patient with a complaint about his or her voice consists of an extensive history including medications, previous surgeries, and current and previously diagnosed diseases. Acoustic, perceptual, and physiological assessment of vocal function may reveal evidence of tremor, vocal volume deficiencies, and/or vocal fatigue. Examination of the larynx and vocal folds via flexible endoscopy as well as strobolaryngoscopy is essential to reveal vocal use patterns, asymmetrical vibration, scarring, tremor (of the larynx or other structures), atrophy, or lesions. In the absence of suspected malignancies or frank aspiration due to lack of glottic closure, voice therapy is the treatment of choice for most elderly patients with neurological age-related dysphonias.

**Treatments for Neurological Aging-Related Voice Disorders**

Treatments for elderly patients with mobile vocal folds presenting with dysphonia include behavioral, pharmacological, and surgical approaches. A review of surgical treatments can be found in Ford (1986), Koufman (2000), Postma (1998), and Durson (1996). The use of medications and their relationship to vocal production and vocal aging can be found in the work of Sataloff and colleagues (1997) and Vogel (1995).

Voice therapy for neurological aging-related voice disorders varies, depending on the patient’s complaints, diagnosis, and vocal use requirements. The most common needs of patients with neurological age-related voice disorders are to increase loudness and endurance, to reduce hoarseness or breathy voice qualities, and to maintain a broad pitch range for singing. These needs are met with vocal education including awareness of vocal hygiene; direct vocal exercises; and management of the vocal environment. Prior to voice therapy, and as

### Table 1. Diagnoses of Subjects Age 65 and Older Seen at the University of Pittsburgh Voice Center

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocal fold atrophy</td>
<td>46</td>
<td>23</td>
</tr>
<tr>
<td>Vocal fold paralysis</td>
<td>39</td>
<td>19</td>
</tr>
<tr>
<td>Laryngopharyngeal reflux</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>Essential tremor</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Other neurological disorders</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Muscular tension dysphonia—primary</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Muscular tension dysphonia—secondary</td>
<td>38</td>
<td>19</td>
</tr>
<tr>
<td>Edema</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Spasmodic dysphonia</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

N = 205. The total number of diagnoses is larger since some patients had more than one diagnosis.

### Table 2. The Most Common Voice Symptoms Reported by Patients 65 Years of Age and Older

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of volume</td>
<td>28</td>
</tr>
<tr>
<td>Raspy or hoarse voice</td>
<td>24</td>
</tr>
<tr>
<td>Vocal fatigue</td>
<td>22</td>
</tr>
<tr>
<td>Difficulty breathing during speech</td>
<td>18</td>
</tr>
<tr>
<td>Talking in noisy environments</td>
<td>15</td>
</tr>
<tr>
<td>Loss of clarity</td>
<td>16</td>
</tr>
<tr>
<td>Tremor</td>
<td>7</td>
</tr>
<tr>
<td>Intermittent voice loss</td>
<td>6</td>
</tr>
<tr>
<td>Articulation-related problems</td>
<td>5</td>
</tr>
</tbody>
</table>

Total exceeds 100% as some individuals reported more than one complaint.
part of the diagnostic process, a thorough audiological assessment of the patient should be done. If the patient wears a hearing aid or aids, he or she should wear them for the therapy sessions.

**Vocal Education**

Vocal education coupled with vocal hygiene provides the patient with an understanding of the aging process as it relates to voice use. An understanding of how all body organ systems are affected by normal aging helps to explain why the voice may not have the same quality, pitch range, endurance, or loudness that was present in earlier years. Since the voice is the product of respiratory and vocal tract functions, all of the systems that contribute to the aging of these organs are responsible for the final vocal output. Recently, Murry and Rosen published a vocal education and hygiene program for patients (Murry and Rosen, 2000). This program is an excellent guide for all aging patients with neurological voice disorders. Nursing homes, senior citizen residences, and geriatric specialists should consider offering this outline as the first step in patient education when patients complain of voice disorders.

**Direct Voice Therapy**

Voice therapy is one treatment modality for almost all types of neurological aging-related voice disorders. The recent explosion of knowledge about the larynx is matched by an equal growth of interest in its physiology, its disorders, and their treatment. Increased use of laryngeal imaging and knowledge of laryngeal physiology have provided a base for behavioral therapy that is increasingly focused on the specific nature of the observed pathophysiology. While treatment is designed to restore maximum vocal function, the aging process of weakness, muscle wasting, and system endurance may not restore the voice to its youthful characteristics. Rather, the desired goals should be effective vocal communication and forestalling continued vocal deterioration (Sataloff et al., 1997).

Specific techniques for the aging voice have evolved from our understanding of the aging neuromuscular process. Confident voice is the voice that one might typically use to describe or discuss confidential matters. Theoretically, it is produced with minimal vocal fold contact. The confidential voice technique is used to (1) eliminate hyperfunctional and traumatic behaviors; (2) allow lesions such as vocal nodules to heal in the absence of continued pounding; (3) eliminate excessive muscular tension and vocal fatigue; (4) reset the internal volume meter; and (5) force a heightened awareness of voice use and the vocal environment. The goal is to create healthier vocal folds and a neutral state from which healthy voice use can be taught and developed through a variety of other techniques (Verdolini-Marston, Burke, and Lessass, 1995; Leddy, Samlan, and Poburka, 1997).

The confidential voice technique is appropriately used to treat benign lesions, muscle tension dysphonia, hyperfunctional dysphonia, and vocal fatigue in the early postoperative period. It is not appropriate for treatment of vocal fold paralysis, conditions with incomplete glottal closure, or a scarred vocal fold.

Resonant voice, or voice with forward focus, usually refers to an easy voice associated with vibratory sensations in facial bones (Verdolini-Marston, Burke, and Lessass, 1995). Therapy focuses on the production of this voice primarily through feeling and hearing. Exercises to place the vocal mechanism in a specific manner coupled with humming help the patient identify to optimum pitch/placement for maximum voice quality. Resonant voice therapy is described as being useful in the treatment of vocal fold lesions, functional voice problems, mild vocal atrophy, and paralysis.

Manual circumlaryngeal massage (manual laryngeal musculoskeletal tension reduction) is a direct, hands-on approach in which the clinician massages and manipulates the laryngeal area in a particular manner while observing changes in voice quality as the patient phonates. The technique was first proposed by Aronson (1990) and later elaborated by Morrison and Rammage (1993). Roy and colleagues reported on their use of the massage technique in controlled studies (Roy, Bless, and Heisey, 1997). They reported almost normal voice following a single session in 93% of 17 subjects with hyperfunctional dysphonia.

General body massage in which muscles are kneaded and manipulated is known to reduce muscle tensions. This concept is adapted to massage the muscles in the laryngeal area. One focus of the circumlaryngeal massage is to relieve the contraction of those muscles and allow the larynx to lower. This technique is most often used with patients who report neck or upper body tension or stiffness, tenderness in the neck muscles, odynophonia, or those who demonstrate rigid postures. Vocal function exercises are designed to pinpoint and exercise specific laryngeal muscles. The four steps address warm-up of the muscles, stretching and contracting of muscles, and building muscle power. The softness of the productions is said to increase muscular and respiratory effort and control. The exercises are hypothesized to restrengthen and balance laryngeal musculature, to improve vocal fold flexibility and movement, and to rebalance airflow (Stemple, Glaze, and Gerdeman, 1995).

Ramig and colleagues developed a structured intensive therapy program, the Lee Silverman voice treatment program, of four sessions per week for 4 weeks specifically for patients with idiopathic Parkinson’s disease (Ramig, Bonitati, and Lemke, 1994). Since then, the efficacy of the treatment for this population has been extended to include aging patients and patients with other forms of progressive neurological disease. This treatment method of voice therapy may be the most promising of all for neurological aging-related voice disorders.

The Lee Silverman voice treatment program is based on the principle that, to counteract the physical effects of reduced amplitude of motor acts including voice and speech production, rigidity, bradykinesia, and reduction in respiratory effort, it is necessary to push the entire
phonatory mechanism to exert greater effort by focusing on loudness. To increase loudness, the respiratory system must provide more driving power, and the vocal folds must adduct more completely. Indeed, it seems that the respiratory, the laryngeal, and the articulatory mechanisms all benefit from the effort to increase loudness.

The program is highly structured and involves five essential concepts: (1) voice is the focus; (2) a high degree of effort is required; (3) treatment is intensive; (4) the patient’s self-perception must be calibrated; and (5) productions and outcomes are quantified. The scope of this article does not permit inclusion of the extensive literature available on this method or an extensive description of the therapy protocol, which is available in published form.

The accent method, originally developed by Smith, has been used to treat all types of dysphonias (Smith and Thyme, 1976). It has been adopted more widely abroad than in the United States and focuses on breathing as the underlying control mechanism of vocal output and uses accentuated and rhythmic movements of the body and then of voicing. Easy voice production with an open-throat feeling is stressed, and attention is paid primarily to an abdominal/diaphragmatic breathing pattern. This method is useful for treating those individuals with vocal fatigue, endurance problems, or overall volume weakness.

All voice therapy is a directed way of changing a particular behavior or set of behaviors. Regardless of the methods used, voice therapy demands the cooperation of the patient in ways that may be novel and unusual. Voice therapy differs from the medical approach, which requires only that a pill be taken or an injection received. It differs from the surgical approach, wherein the surgeon does the work. It differs from the work of voice and acting coaches, who work to enhance and strengthen a normal voice. Voice clinicians work with individuals who never thought about the voice until they acquired a voice disorder. They are primarily interested in rapid restoration of normal voice, a task that cannot always be accomplished.

**Vocal Tremor**

One neurological aging-related disorder that often resists change is vocal tremor. Tremor often accompanies many voice disorders having a neurological component. Vocal tremor has been treated in the past with medications, and in some cases with laryngeal framework surgery, when vocal fold atrophy is also diagnosed. The specific therapeutic techniques presented in this article may also be helpful in reducing the perception of tremor especially those that focus on increasing vocal fold closure (i.e., Lee Silverman voice treatment and the Accent Method).

Finally, the treatment of aging-related dysphonias should include techniques used in training singers and actors (Sataloff et al., 1997). General physical conditioning, warmup, increased respiratory function exercises, and the ability to monitor voice change help to maintain voice or retard vocal weakness, fatigue, and loss of clear voice quality.

The vocal environment should not be ignored as a factor in communication for patients with neurological aging-related voice disorders. Voice use is maximized in environments where background noise is minimal, sound absorption materials such as rugs and cushions are used in large meeting rooms, and proper lighting is available to help with visual components of communication.

**Summary**

Progressive neurological aging-related disorders offer a challenge to the speech-language pathologist. In the absence of surgery for vocal fold paralysis, the patient with mobile vocal folds and a neurological disorder may benefit from specific exercises to maintain vocal communication. Diagnosis, which identifies the vocal use habits of the patient is critical to identify strategies and the specific exercises needed to maintain and/or improve voice production.

*See also VOICE DISORDERS OF AGING.*

—Thomas Murry

**References**


A professional voice user is a person whose job function critically depends on use of the voice. Not only singers and actors but teachers, lawyers, clergy, counselors, air traffic controllers, telemarketers, firefighters, police, and auctioneers are among those who use their voices significantly in their line of work.

Probably the preponderance of professional voice users who seek treatment for voice problems have voice-induced conditions. Typically, such conditions involve either phonotrauma or functional problems. The full range of non-use-related vocal pathologies may occur in professional voice users as well, at about the same rate as in the population at large. However, special considerations may be required in therapy for professional voice users because of their job demands.

The teaching profession is at highest risk for voice problems. In 1999, teachers made up between 5% and 6% of the employed population in the United States. At any given time, between one-fifth and one-half of teachers in the United States and elsewhere are experiencing a voice problem (Sapir et al., 1993; Russell, Oates, and Greenwood, 1998; E. Smith et al., 1998). Voice problems appear to occur at about the same rates among singers. Other occupations at risk for voice problems are lawyers, clergy, telemarketers, and possibly even counselors and social workers. Increasingly, phonotrauma is considered an occupational hazard in these populations (Villkman, 2000).

A new occupational hazard for voice problems has recently surfaced in the form of repetitive strain injury. This condition, one of the fastest growing occupational injuries in the United States in general, involves weakness and pain from somatic overuse. Symptoms of repetitive strain injury typically begin in the fingers after keyboard use. However, laryngeal symptoms may develop if the individual replaces the keyboard with voice keyboard use. Not only singers and actors but teachers, lawyers, clergy, counselors, air traffic controllers, telemarketers, firefighters, police, and auctioneers are among those who use their voices significantly in their line of work.

The goal of treatment for professional voice users is to restore the best possible voice use—and, where relevant, anatomy and physical function—relative to the job.
in question. Vocal hygiene, including hydration and reflux control, plays a role in most treatment programs for professional voice users (see VOCAL HYGIENE). Surgical management may be appropriate in selected cases. However, for most professional voice users, the mainstay of intervention for voice problems is behavioral work on voice production, or voice therapy.

Traditional therapy for phonotrauma in professional groups that use the voice quantitatively (e.g., teachers, clergy, attorneys), that is, over an extended period of time or at sustained loudness, has emphasized voice conservation. In this approach, however, individuals are limited at least as much by the treatment as by the disease. The current emphasis is on training individuals to meet their voice needs while they recover from existing problems, and to prevent new ones. An intermediate vocal fold configuration, involving slight separation of the vocal processes during phonation, appears relevant to this goal (Berry et al., 2001). A variety of training methods are available for this approach to vocalization, including Lessac-Madsen resonant voice therapy (Verdolini, 2000), Vocal Function Exercises (Stemple et al., 1994), the Accent Method (S. Smith and Thyme, 1976), and flow mode therapy (see, e.g., Gauflin and Sundberg, 1989). Training in this general laryngeal configuration appears to be more effective than vocal hygiene intervention alone and more effective than intensive respiratory training in reducing self-reported functional problems due to voice in at least one class of professional voice users, teachers (Roy et al., 2001).

Therapy for individuals with qualitative both qualitative voice needs recognizes that a special sound of the voice is required occupationally. Therapy for performers—singers and actors—with voice problems is conceptually challenging, for many reasons. Vocal performers have exacting voice needs, which may be complicated by pathology; the voice training of singers and actors is not standardized; few scientific studies on training efficacy exist; and performers are subject to a suite of special personality, career, and lifestyle issues. All of these factors make many speech-language practitioners feel that a specialty focus on vocology is important in working with performing artists.

Voice therapy for performers often replicates voice pedagogy methods. The primary differences are an emphasis on injury reduction and a shorter-term intervention, with specific, measurable goals, in voice therapy. The most comprehensive technical framework for professional voice training in general has been proposed by Estill (2000). The system identifies 11 or 12 physical “degrees of freedom,” such as voice onset type, false vocal fold position, laryngeal height, palatal position, and aryepiglottic space, that are independently varied to create “recipes” for a variety of sung and spoken voice qualities. Research conducted thus far has corroborated some aspects of the approach (e.g., Titze, 2002). The system recently has gained currency in voice therapy as well as vocal pedagogy. Voice training for acting tends to be less technically oriented and more “meaning driven” than singing training (e.g. Linklater, 1997). However, exceptions exist. Also, theatre and increasingly singing training and voice therapy incorporate general body work (alignment, movement) as a central part of training.

In respect to training modalities, traditional speech-language pathology models tend to be more analytical and less experiential than typical performing arts models of training. The motor learning literature indicates that the performers may be right. The literature describes a critical dependence of motor learning on sensory processing and deemphasizes mechanical instruction (Verdolini, 1997; see also Wulf, Höß, and Prinz, 1998). The motor learning literature also clearly indicates the need for special attention to transfer in training. Skills acquired in a clinic or studio may transfer poorly to untrained stimuli in untrained environments if less specific transfer exercises are used. Biofeedback may be a useful adjunct to voice therapy and training; however, cautions exist. Terminal biofeedback, provided after the completion of performance, contributes to greater learning than on-line feedback, which occurs during ongoing performance (Armstrong, 1970, cited in Schmidt and Lee, 1999, pp. 316–317). Also, systematic fading of biofeedback support appears critical for transfer.

The voice therapist may need to address special challenges in the physical and political environments of performers. Stage environments can be frankly toxic, and compromising to vocal and overall physical health. Specific noxious substances that have been measured on stages include aromatic diisocyanates, Penicillium frequentans and formaldehyde in cork granulate, cobalt and aluminum (pigment components), and alveolar-size quartz sand (Richter et al., 2002). Open-air performing environments can present particular vocal challenges to performers, especially if these are unmiked. Heavy costumes weighing 80–90 lb or more and unusual, contorted postures required during vocal performance may add further challenges and may even contribute to injury.

Politically, performers may find themselves contractually linked to heavy performance schedules without the possibility of rest if they are ill or vocally indisposed. Performers are threatened with loss of income, loss of health care benefits, and loss of professional reputation if they refuse to perform when they should not. Another political issue has to do with directors’ drive toward meeting commercial goals. Such goals may dictate vocal practices that are at odds with performers’ best interest. Directors and producers may sometimes show little concern for performers’ vocal health, because numerous vocalists are available to replace injured ones who are unwilling or unable to perform.

It is probably safe to say that individuals who are drawn to vocal performance are more extroverted, and more emotionally variable, on average, than many individuals in the population at large. The vocal practitioner should be comfortable dealing with performers’ individual personal styles. Moreover, mental attitude toward performance plays a central role in the performing domain. The principles of sports psychology fully apply to
the performing arts. A robust finding is that intermediate anxiety levels, as opposed to low or high anxiety, tend to maximize physical performance. Performers need to find ways to establish intermediate arousal states and stay there even in high-stress situations. Also, the direction of attention appears key for distinguishing “chokers” (people who tend to perform poorly under pressure) from persons who perform well under high stress. According to some reports, chokers tend to show a predominance of left hemisphere activation when under the gun, implying verbal analytic thinking and evaluative self-awareness. High-level performers tend to show more distributed brain activation, including right-hemisphere activity consistent with imagery and target awareness (Crews, 2001). Many other findings from the sports psychology literature are applicable to attitude issues in vocal performance.

Vocal performers may have erratic lifestyles that are linked to their jobs. Touring groups literally may live on buses. Exercise and fresh air may be restricted. Daily routines may be nonexistent. Pay may be poor and sporadic. Benefits often are not provided unless the performers belong to a union. Vocal performers with voice problems often cannot pay for treatment because their voice problems lead to lack of employment and thus lack of income and benefits. Clinics wishing to work with professional voice users should be equipped to provide some form of fiscal support for treatment.

Practitioners working with vocal performers agree that no single individual can fully assist a vocalist with voice problems. Rather, convergent efforts are required across specialties, to minimally include an otolaryngologist, speech-language pathologist, voice teacher or coach, and, patient. Different individuals take the lead, depending on the issues at hand. The physician is responsible for medical issues. The speech-language pathologist and voice teacher generally work together on technical issues. The voice teacher is the most appropriate person to address career issues with the performer, particularly issues that bear on a potential mismatch between the individual’s aspirations and capabilities. The importance of communication across individuals within the team cannot be overemphasized.

—Katherine Verdolini

References


Further Readings


Part II: Speech
Apraxia of speech is a phonetic-motoric disorder of speech production caused by inefficiencies in the translation of a well-formed and filled phonologic frame to previously learned kinematic parameters assembled for carrying out the intended movement, resulting in intra- and inter-articulator temporal and spatial segmental and prosodic distortions. It is characterized by distortions of segments, intersegment transitionalization resulting in extended durations of consonants, vowels and time between sounds, syllables and words. These distortions are often perceived as sound substitutions and the misassignment of stress and other phrasal and sentence-level prosodic abnormalities. Errors are relatively consistent in location within the utterance and invariable in type. It is not attributable to deficits of muscle tone or reflexes, nor to deficits in the processing of auditory, tactile, kinesthetic, proprioceptive, or language information. (McNeil, Robin, and Schmidt, 1997, p. 329)

The kernel perceptual behaviors that differentiate apraxia of speech (AOS) from other motor speech disorders and from phonological paraphasia are (1) lengthened segment (slow movements) and intersegment (segment segregation) durations (overall slowed speech), resulting in (2) abnormal prosody across multisyllable words and phrases, with a tendency to make errors on more stressed than unstressed syllables; (3) relatively consistent trial-to-trial location of errors and relatively nonvariable error types; and (4) impaired measures of coarticulation. Although apraxic speakers may produce a preponderance of sound substitutions, these substitutions do not serve as evidence of either AOS or phonemic paraphasia. Sound distortions serve as evidence of a motor-level mechanism or influence in the absence of an anatomical explanation; however, they are not localizable to one part of the motor control architecture and, taken alone, do not differentiate AOS from the dysarthrias. Acoustically well-produced (non-distorted) sound-level serial order (e.g., perseverative, anticipatory, and exchange) errors that cross word boundaries are not compatible with motor planning- or programming-generated mechanisms and are attributable to the phonological encoding mechanism.

Although this motor speech disorder has a languorous and tortuous theoretical and clinical history and is frequently confused with other motor speech disorders and with phonemic paraphasia, a first-pass estimate of some of its epidemiological characteristics has been presented by McNeil, Doyle, and Wambaugh (2000).

Based on retrospective analysis of the records of 3417 individuals evaluated at the Mayo Clinic for acquired neurogenic communication disorders, including dysarthria, AOS, aphasia, and other neurogenic speech, language, and cognitive disorders, Duffy (1995) reported a 4.6% prevalence of AOS. Based on this same retrospective analysis of 107 patient records indicating a diagnosis of AOS, Duffy reported that 58% had a vascular etiology and 6% presented with a neoplasm. One percent presented with a seizure disorder and 16% had a diagnosis of degenerative disease, including Creutzfeldt-Jakob disease and leukoencephalopathy (of the remaining, 9% were unspecified, 4% were associated with dementia, and 3% were associated with primary progressive aphasia). In 15% of cases the AOS was traumatically induced (12% neurosurgically and 3% concomitant with closed head injury), and in the remaining cases the cause was undetermined or was of mixed etiology. Without doubt, these proportions are influenced by the type of patients typically seen at the Mayo Clinic, and may not be representative of other patient care sites.

Among all of the acquired speech and language pathologies of neurological origin, AOS may be the most infrequent. Its occurrence unaccompanied by dysarthria, aphasia, limb apraxia, or oral-nonspeech apraxia is extremely rare. Comorbidity estimates averaged across studies and summarized by McNeil, Doyle, and Wambaugh (2000) indicated an AOS/oral-nonspeech apraxia comorbidity of 68%, an AOS/limb apraxia comorbidity of 67%, an AOS/limb apraxia and oral-nonspeech apraxia comorbidity of 83%, an AOS/aphasia comorbidity of 81%, and an AOS/dysarthria comorbidity of 31%. Its frequent co-occurrence with other disorders and its frequent diagnostic confusion with those disorders that share surface features with it suggest that the occurrence of AOS in isolation (pure AOS) is extremely rare.

The lesion responsible for AOS has been studied since Darley (1968) and Darley, Aronson, and Brown (1975) proposed it as a neurogenic speech pathology that is theoretically and clinically different from aphasia and the dysarthrias. Because Darley defined AOS as a disorder of motor programming, the responsible lesion has been sought in the motor circuitry, especially in Broca’s area. Luria (1966) proposed that the frontal lobe mechanisms for storing and accessing motor plans or programs for limb gestures or for speech segments were represented in Broca’s area. He also proposed the facial region of the postcentral gyrus in the parietal lobe as a critical area governing coordinated movement between gestures (speech or nonspeech). AOS-producing lesions subtending Broca’s area (Mohr et al., 1978) as well as those in the postcentral gyrus (Square, Darley, and Sommers, 1982; Marquardt and Sussman, 1984; McNeil et al., 1990) have received support. Retrospective studies of admittedly poorly defined and poorly described persons purported to have AOS (e.g., Kertesz, 1984) do not show a single site or common cluster of lesion sites responsible for the disorder. Prospective studies of the AOS-producing lesion have been undertaken by a number of investigators. Deutsch (1984) was perhaps the first to conduct a prospective search, with a result that set the stage for most of the rest of the results to follow. He found that 50% of his AOS subjects (N = 18) had a lesion in the frontal lobe and 50% had posterior lesions. Marquardt and Sussman’s (1984) prospective study of 12 subjects with AOS also failed to reveal a consistent relationship among lesion location (cortical...
versus subcortical, anterior versus posterior), lesion volume, and the presence or absence of AOS. Dronkers (1997) reported that 100% of 25 individuals with AOS had a discrete left hemispheric cortical lesion in the pre-central gyrus of the insula. One hundred percent of a control group of 19 individuals with left hemispheric lesions in the same arterial distribution as the AOS subjects but without the presence of AOS were reported to have had a complete sparing of this specific region of the insula. McNeil et al. (1990) reported computed tomographic lesion data from four individuals with AOS unaccompanied by other neurogenic speech or language pathologies. The only common lesion site for these “pure” apraxic speakers was in the facial region of the left postcentral gyrus. Two of the four AOS subjects had involvement of the insula, while two of three subjects with phonemic paraphasia (diagnosed with conduction aphasia) had a lesion in the insula. Two of the four subjects with AOS and one of the three subjects with conduction aphasia evinced involvement of Broca’s area. The unambiguous results of the Dronkers study notwithstanding, the lesions responsible for AOS remain open to study. It is clear, however, that the major anterior/posterior divisions common to aphasiology and traditional neurology as sites responsible for nonfluent/ fluent (respectively) disorders of speech production are challenged by the AOS/lesion data that are available to date.

**Theoretical Accounts**

The study of and clinical approach to AOS operate under a scientific paradigm generally consistent with the mechanisms ascribed to apraxia. That is, the majority of practitioners view AOS as a disorder of previously learned movements that is different from other speech movement disorders (i.e., the dysarthrias). The diagnosis can be confidently applied when assurance can be obtained that the person has the cognitive or linguistic knowledge underlying the intended movement and the fundamental structural and sensorimotor abilities to carry out the movement. Additionally, most definitions of apraxia suggest an impairment of movements carried out volitionally but executed successfully when performed automatically. The diagnosis requires that patients display the ability to process the language underlying the movement. These criteria are generally consistent with those used for the identification of other apraxias, including oral nonspeech (buccofacial), writing (agraphic), and limb apraxia.

Although AOS is predominantly viewed as a disorder of motor programming (Wertz, LaPointe, and Rosenbek, 1984), derived from its historical roots based in other apraxias (particularly limb-kinetic apraxia; McNeil, Doyle, and Wambaugh, 2000), there are competing theories. Whiteside and Varley (1998) proposed a deficit of the direct phonetic encoding route to account for AOS. In this theory, normal speech production involves the retrieval from storage of verbal motor patterns for frequently used syllables (the direct route), or the patterns are calculated anew (presumably from smaller verbal motor patterns) by an indirect route. Speech produced by normal speakers for infrequently occurring syllables, using the indirect route, are said to share many of the core features of apraxic speakers, such as (1) articulatory prolongation, (2) syllable segregation, (3) inability to increase the speech rate and maintain articulatory integrity, and (4) reduced coarticulation. AOS is therefore proposed to be a deficit of the direct encoding route, with a reliance on the indirect encoding route.

Based on experimental evidence that the phonological similarity effect should not be present in persons with AOS, Rogers and Storkel (1998, 1999) hypothesized a reduced buffer capacity as the mechanism responsible for AOS. In this account, the apraxic speaker with a reduced buffer capacity is required to reload or reprogram the appropriate (unspecified) buffer in a feature-by-feature, sound-by-sound, syllable-by-syllable, or motor-control-variable-by-motor-control-variable fashion. This requirement would give rise to essentially the same observable features of AOS as those commonly used to define the entity and consistent with the observable features discussed earlier.

Van der Merwe (1997) proposed a model of sensorimotor speech disorders in which AOS is defined as a disorder of motor planning. Critical to this view is the separation of motor plans from motor programs. In this model, motor plans carry information (e.g., lip rounding, jaw depression, glottal closure, raising or lowering of the tongue tip, interarticulator phasing/coarticulation) that is articulator-specific, not muscle-specific. Motor plans are derived from specific speech sounds and specify the spatial and temporal goals of the planned unit. Motor programs, on the other hand, specify the movement parameters (e.g., muscle tone, direction, force, range, and rate of movement) of specific muscles or muscle groups. For Van der Merwe, disorders of motor programs result in the dysarthrias and cannot account for the different set of physiological and behavioral signs of AOS. The attributes ascribed to motor plans in this model are consistent with the array of cardinal behavioral features of AOS.

Though their view is expanded from the traditional view of AOS as simply a disorder of motor programming, McNeil, Doyle, and Wambaugh (2000) argue that a combined motor planning and motor programming impairment as specified by Van der Merwe (1997) is required to account for the array of well-established perceptual, acoustic, and physiological features.

Other theoretical accounts of AOS include the overspecification of phonological representations theory of Dogil, Mayer, and Vollmer (1994) and the coalitional/dynamical systems breakdown theory of Kelso and Tuller (1981). These accounts have received considerably less examination in the literature and will not be described here.

AOS is an infrequently occurring pathology that is clinically recognized by most professionals dedicated to the management of speech production disorders. It is classified as a motor speech disorder in the scientific literature. When it occurs, it is frequently accompanied
by other speech, language, and apraxic disorders. Its defining features are not widely agreed upon; however, evidence from perceptual, acoustic, kinematic, aerodynamic, and electromyographic studies, informed by recent models of speech motor control and phonological encoding, have led to clearer criteria for subject/patient selection and a resurgence of interest in its proposed mechanisms. The lesions responsible for acquired AOS remain a matter for future study.

See also APRAXIA OF SPEECH: TREATMENT; DEVELOPMENTAL APRAXIA OF SPEECH.

—Malcolm R. McNeil and Patrick J. Doyle

References


Further Readings


Apraxia of Speech: Treatment

In the years since Darley (1968) first described apraxia of speech (AOS) as an articulatory programming disorder that could not be accounted for by disrupted linguistic or fundamental motor processes, considerable work has been done to elucidate the perceptual, acoustic, kinematic, aerodynamic, and electromyographic features that characterize AOS (cf. McNeil, Robin, and Schmidt, 1997; McNeil, Doyle, and Wambaugh, 2000). Explanatory models consistent with these observations have been proposed (Van der Merwe, 1997).

Overwhelmingly, the evidence supports a conceptualization of AOS as a neurogenic speech disorder caused by inefficiencies in the specification of intended articulatory movement parameters or motor programs which result in intra- and interarticulator temporal and spatial segmental and prosodic distortions. Such movement distortions are realized as extended segmental, intersegmental transitionalization, syllable and word durations, and are frequently perceived as sound substitutions, the misassignment of stress, and other phrasal and sentence-level prosodic abnormalities. (McNeil, Robin, and Schmidt, 1997, p. 329)

Traditional and contemporary conceptualizations of the disorder have resulted in specific assumptions regarding appropriate tactics and targets of intervention, and a number of treatment approaches have been proposed that seek to enhance (1) postural shaping and phasing of the articulators at the segmental and syllable levels, and (2) segmental sequencing of longer speech units (Square-Storer and Hayden, 1989).

More recently, arguments supporting the application of motor learning principles (Schmidt, 1988) for the purposes of specifying the structure of AOS treatment sessions have been proposed, based on evidence that such principles facilitate learning and retention of motor routines involved in skilled limb movements (Schmidt, 1991). The empirical support for each approach to treatment is reviewed here.

Enhancing Articulatory Kinematics at the Segmental Level. Several facilitative techniques have been recommended to enhance postural shaping and phasing of the articulators at the segmental and syllable levels and have been described in detail by Wertz, LaPointe, and Rosenbek (1984). These techniques include (1) phonetic derivation, which refers to the shaping of speech sounds based on corresponding nonspeech postures, (2) progressive approximation, which involves the gradual shaping of targeted speech segments from other speech segments, (3) integral stimulation and phonetic placement, which employ visual models, verbal descriptions and physical manipulations to achieve the desired articulatory posture, and movement, and (4) minimal pairs contrasts, which requires patients to produce syllable or word pairs in which one member of the pair differs minimally with respect to manner, place, or voicing features from the other member of the pair.

Several early studies examined, in isolation or in various combinations, the effects of these facilitative techniques on speech production, and reported positive treatment responses (Rosenbek et al., 1973; Holtzapple and Marshall, 1977; Deal and Florance, 1978; Thompson and Young, 1983; LaPointe, 1984; Wertz, 1984). However, most of these studies suffered from methodological limitations, including inadequate subject selection criteria, nonreplicable treatment protocols, and pre-experimental research designs, which precluded firm conclusions regarding the validity and generalizability of the reported treatment effects. Contemporary investigations have addressed these methodological shortcomings and support earlier findings regarding the positive effects of treatment techniques aimed at enhancing articulatory kinematic aspects of speech at the sound, syllable, and word levels.

Specifically, in a series of investigations using single-subject experimental designs Wambaugh and colleagues examined the effects of a procedurally explicit treatment protocol employing the facilitative techniques of integral stimulation, phonetic placement, and minimal pair contrasts in 11 well-described subjects with AOS (Wambaugh et al., 1996, 1998, 1999; Wambaugh, West, and Doyle, 1998; Wambaugh and Cort, 1998; Wambaugh, 2000). These studies revealed positive treatment effects on targeted phonemes in trained and untrained words for all subjects across all studies, and positive maintenance effects of targeted sounds at 6 weeks post-treatment. In addition, two subjects showed positive...
generalization of trained sounds to novel stimulus contexts (i.e., untrained phrases), and one subject showed positive generalization to untrained sounds within the same sound class (voiced stops). These results provide initial experimental evidence that treatment strategies designed to enhance postural shaping and phasing of the articulators are efficacious in improving sound production of treated and untreated words. Further, there is limited evidence that for some patients and some sounds, generalization to untrained contexts may be expected.

**Enhancing Segmental Sequencing of Longer Speech Units.** Several facilitative techniques have been recommended to improve speech production in persons with AOS, based on the premise that the sequencing and coordination of movement parameters required for the production of longer speech units (and other complex motor behaviors) are governed by internal oscillatory mechanisms (Gracco, 1990) and temporal constraints (Kent and Adams, 1989). Treatment programs and tactics grounded in this framework employ techniques designed to reduce or control speech rate while enhancing the natural rhythm and stress contours of the targeted speech unit. The effects of several such specific facilitative techniques have been studied. These include metronomic pacing (Shane and Darley, 1978; Dworkin, Abkarian, and Johns, 1988; Dworkin and Abkarian, 1996; Wambaugh and Martinez, 1999), prolonged speech (Southwood, 1987), vibrotactile stimulation (Rubow et al., 1982), and intersystemic facilitation (i.e., finger counting) (Simmons, 1978). In addition, the effects of similarly motivated treatment programs, melodic intonation therapy (Sparks, 2001) and surface prompts (Square, Chumpelik, and Adams, 1985), have also been reported.

As with studies examining the effects of techniques designed to enhance articulatory kinematic aspects of speech at the segmental level, the empirical evidence supporting the facilitative effects of rhythmic pacing, rate control, and stress manipulations on the production of longer speech units in adults with AOS is limited. That is, among the reports cited, only five subjects were studied under conditions that permit valid conclusions to be drawn regarding the relationship between application of the facilitative technique and the dependent measures reported (Southwood, 1987; Dworkin et al., 1988; Dworkin and Abkarian, 1996; Wambaugh et al., 1999). Whereas each of these studies reported positive results, it is difficult to compare them because of differences in the severity of the disorder, in the frequency, duration, and context in which the various facilitative techniques were applied, in the behaviors targeted for intervention, and in the extent to which important aspects of treatment effectiveness (i.e., generalized effects) were evaluated. As such, the limited available evidence suggests that techniques that reduce the rate of articulatory movements and highlight rhythmic and prosodic aspects of speech production may be efficacious in improving segmental coordination in longer speech units. However, until these findings can be systematically replicated, their generalizability remains unknown.

**General Principles of Motor Learning.** The contemporary explication of AOS as a disorder of motor planning and programming has given rise to a call for the application of motor learning principles in the treatment of AOS (McNeil et al., 1997, 2000; Ballard, 2001). The habituation, transfer, and retention of skilled movements (i.e., motor learning) and their controlling variables have been studied extensively in limb systems from the perspective of schema theory (Schmidt, 1975). This research has led to the specification of several principles regarding the structure of practice and feedback that were found to enhance retention of skilled limb movements post-treatment, and greater transfer of treatment effects to novel movements (Schmidt, 1991). Three such principles are particularly relevant to the treatment of AOS: (1) the need for intensive and repeated practice of the targeted skilled movements, (2) the order in which targeted movements are practiced, and (3) the nature and schedule of feedback.

With respect to the first of these principles, clinical management of AOS has long espoused intensive drill of targeted speech behaviors (Rosenbek, 1978; Wertz et al., 1984). However, no studies have examined the effects of manipulating the number of treatment trials on the acquisition and retention of speech targets in AOS, and little attention has been paid to the structure of drills used in treatment. That is, research on motor learning in limb systems has shown that practicing several different skilled actions in random order within training sessions facilitates greater retention and transfer of targeted actions than does blocked practice of skilled movements (Schmidt, 1991). This finding has been replicated by Knock et al. (2000) in two adult subjects with AOS in the only study to date to experimentally manipulate random versus blocked practice to examine acquisition, retention, and transfer of speech movements.

The final principle to be discussed concerns the nature and schedule of feedback employed in the training of skilled movements. Two types of feedback have been studied, knowledge of results (KR) and knowledge of performance (KP). KR provides information only with respect to whether the intended movement was performed accurately or not. KP provides information regarding aspects of the movement that deviate from the intended action and how the intended action is to be performed. Schmidt and Lee (1999) argue that KP is most beneficial during the early stages of training but that KR administered at low response frequencies promotes greater retention of skilled movements. Both types of feedback are frequently employed in the treatment of AOS. Indeed, the facilitative techniques of integral stimulation and phonetic placement provide the type of information that is consistent with the concept of KP. However, these facilitative techniques are most frequently used as antecedent conditions to enhance target performance, and response-contingent feedback frequently takes the form of KR. The effects of the nature,
schedule, and timing of performance feedback have not been systematically investigated in AOS.

In summary, AOS is a treatable disorder of motor planning and programming. Studies examining the effects of facilitative techniques aimed at improving postural shaping and phasing of the articulators at the segmental level and sequencing and coordination of segments into long utterances have reported positive outcomes. These studies are in need of carefully controlled systematic replications before generalizability can be inferred. Further, the effects of motor learning principles (Schmidt and Lee, 1999) on the habituation, maintenance, and transfer of speech behaviors require systematic evaluation in persons with AOS.

—Patrick J. Doyle and Malcolm R. McNeil

References


Further Readings


Aprosodia

Prosody consists of alterations in pitch, stress, and duration across words, phrases, and sentences. These same parameters are defined acoustically as fundamental frequency, intensity, and timing. It is the variation in these parameters that not only provides the melodic contour of speech, but also invests spoken language with linguistic and emotional meaning. Prosody is thus crucial to conveying and understanding communicative intent.

The term “aprosodia” was first used by Monrad-Krohn (1947) to describe loss of the prosodic features of speech. It resurfaced in the 1980s in the work of Ross and his colleagues to refer to the attenuated use of and decreased sensitivity to prosodic cues by right hemisphere damaged patients (Ross and Mesulam, 1979; Ross, 1981; Gorelick and Ross, 1987).

Prosodic deficits in expression or comprehension can accompany a variety of cognitive, linguistic, and psychiatric conditions, including dysarthria and other motor...
speech disorders, aphasia, chronic alcoholism, schizophrenia, depression, and mania, as well as right hemisphere damage (RHD) (Duñey, 1995; Myers, 1998; Monnot, Nixon, Lovallo, and Ross, 2001). The term *aprosodia*, however, typically refers to the prosodic impairments that can accompany RHD from stroke, head injury, or progressive neurologic disease with a right hemisphere focus. Even the disturbed prosody of other illnesses, such as schizophrenia, may be the result of alterations in right frontal and extrapyramidal areas, areas considered important to prosodic impairment subsequent to RHD (Sweet, Primeau, Fichtner et al., 1998; Ross et al., 2001).

The clinical presentation of expressive aprosodia is a flattened, monotonic, somewhat robotic, stilted prosodic production characterized by reduced variation in prosodic features and somewhat uniform intersyllable pause time. The condition often, but not always, accompanies flat affect, a more general form of reduced environmental responsivity, reduced sensitivity to the paralinguistic features of communication (gesture, body language, facial expression), and attenuated animation in facial expression subsequent to RHD. Aprosodia can occur in the absence of dysarthria and other motor speech disorders, in the absence of depression or other psychiatric disturbances, and in the absence of motor programming deficits typically associated with apraxia of speech. Because it is associated with damage to the right side of the brain, it usually occurs in the absence of linguistic impairments (Duñey, 1995; Myers, 1998).

Expressive aprosodia is easily recognized in patients with flat affect. Deficits in prosodic perception and comprehension are less apparent in clinical presentation. It is important to note that receptive and expressive prosodic processing can be differentially affected in aprosodia.

First observed in the emotional domain, aprosodia has also been found to occur in the linguistic domain. Thus, patients may have problems both *encoding* and *decoding* the tone of spoken messages and the intention behind the message as conveyed through both linguistic and emotional prosody.

In the acute stage, patients with aprosodia are usually unaware of the problem until it is pointed out to them. Even then, they may deny it, particularly if they suffer from other forms of denial of deficit. Severity of neglect, for example, has been found to correlate with prosodic deficits (Starkstein, Federoff, Price et al., 1994). In rare cases, aprosodia may last for months and even years when other signs of RHD have abated. Patients with persistent aprosodia may be aware of the problem but feel incapable of correcting it.

Treatment of aprosodia is often limited to training patients to adopt compensatory techniques. Patients may be taught to attend more carefully to other forms of emotional expression (e.g., gesture, facial expression) and to signal mood by explicitly stating their mood to the listener. There has, however, been at least one report of successful symptomatic treatment using pitch biofeedback and modeling (Stringer, 1996). Treatment has been somewhat limited by uncertainty about the underlying mechanisms of aprosodia. It is not clear the extent to which expressive aprosodia is a motor problem, a pragmatic problem, a resource allocation problem, or some combination of conditions. Similarly, it is not clear whether receptive aprosodia is due to perceptual interference in decoding prosodic features, to restricted attention (which may reduce sensitivity to prosodic cues), or to some as yet unspecified mechanism.

Much of the research in prosodic processing has been conducted to answer questions about the laterality of brain function. Subjects with unilateral left or right brain damage have been asked to produce linguistic and emotional prosody in spontaneous speech, in imitation, and in reading tasks at the single word, phrase, and sentence level. In receptive tasks they have been asked to determine the emotional valence of expressive speech and to discriminate between various linguistic forms and emotional content in normal and in filtered-speech paradigms. Linguistic tasks include discriminating between nouns and noun phrases based on contrastive stress patterns (e.g., *greenhouse* versus *green house*); using stress patterns to identify sentence meaning (*Joe gave Ella flowers* versus *Joe gave Ella *flowers*); and identifying sentence types based on prosodic contour (e.g., the rising intonation pattern for interrogatives versus the flatter pattern for declaratives).

The emphasis on laterality of function has helped to establish that both hemispheres as well as some subcortical structures contribute to normal prosodic processing. The extensive literature supporting a particular role for the right hemisphere in processing content generated the central hypotheses guiding prosodic laterality research. The first hypothesis suggests that affective or emotional prosody is in the domain of the right hemisphere (Heilman, Scholes, and Watson, 1975; Borod, Koff, Lorch et al., 1985; Blonder, Bowers, and Heilman, 1991). Another hypothesis holds that prosodic cues themselves are lateralized, independent of their function (emotional or linguistic) (Van Lancker and Sidtis, 1992). Finally, lesion localization studies have found that certain subcortical structures, the basal ganglia in particular, play a role in prosodic processing (Cancelliere and Kertesz, 1990; Bravkiv, Dravins, Holtas et al., 1991). Cancelliere and Kertesz (1990) speculated that the basal ganglia may be important not only because of their role in motor control, but also because of their limbic and frontal connections which may influence the expression of emotion in motor action.

Research findings have varied as a function of task type, subject selection criteria, and methods of data analysis. Subjects across and within studies may vary in terms of time post-onset, the presence or absence of neglect and dysarthria, intrahemispheric site of lesion, and severity of attentional and other cognitive deficits. With the exception of site of lesion, these variables have rarely been taken into account in research design. Data analysis has varied across studies. In some studies it has been based on perceptual judgments by one or more listeners, which adds a subjective component. In others, data are
submitted to acoustic analysis, which affords increased objective control but in some cases may not match listener perception of severity of impairment (Ryalls, Joanne, and Feldman, 1987).

In general, acoustic analyses of prosodic productions by RHD patients supports the theory that prosody is lateralized according to individual prosodic cues rather than according to the function prosody serves (emotional versus linguistic). In particular, pitch cues are considered to be in the domain of the right hemisphere. Duration and timing cues are considered to be in the domain of the left hemisphere (Robin, Tranel, and Damasio, 1990; Van Lancker and Sidtis, 1992; Baum and Pell, 1997).

Research suggests that reduced pitch variation and a somewhat restricted pitch range appear to be significant factors in the impaired prosodic production of RHD subjects (Colsher, Cooper, and Graff-Radford, 1987; Behrens, 1989; Baum and Pell, 1997; Pell, 1999a). RHD patients are minimally if at all impaired in the production of emphatic stress. However, they may have an abnormally flat pitch pattern in declarative sentences, less than normal variation in pitch for interrogative sentences, and may produce emotionally toned sentences with less than normal acoustic variation (Behrens, 1988; Emmory, 1987; Pell, 1999). Pitch variation is crucial to signaling emotions, which may explain why impaired production of emotional prosody appears particularly prominent in aprosodia. Interestingly, in the case of tonal languages (e.g., Chinese, Thai, and Norwegian) in which pitch patterns in individual words serve a semantic role, pitch has been found to be a left hemisphere function (Packard, 1986; Ryalls and Reinvang, 1986; Gandour et al., 1992).

Prosodic perception or comprehension deficits associated with aprosodia tend to follow the pattern found in production. Non-temporal properties such as pitch appear to be more problematic than time-dependent properties such as duration and timing (Divenyi and Robinson, 1989; Robin et al., 1990; Van Lancker and Sidtis, 1992). For example, Van Lancker and Sidtis (1992) found that right- and left-hemisphere-damaged patients used different cues to identify emotional stimuli. Patients with RHD tended to base their decisions on durational cues rather than on fundamental frequency variability while left-hemisphere-damaged patients did the opposite. These data suggest a perceptual, rather than a functional (linguistic versus emotional), impairment. Although a study by Pell and Baum (1997) failed to replicate these results, the data are supported by data from dichotic listening and other studies that have investigated temporal versus time-independent cues such as pitch information (Chobor and Brown, 1987; Sidtis and Volpe, 1988; Divenyi and Robinson, 1989; Robin et al., 1990).

Almost all studies of prosodic deficits have focused on whether unilateral brain damage produces prosodic deficits, rather than describing the characteristics of prosodic problems in patients known to have prosodic deficits. The body of laterality research has established that prosodic deficits can occur in both left as well as right hemisphere damage, and has furthered our understanding of the mechanisms and differences in prosodic processing across the hemispheres. However, the focus on laterality has had some drawbacks for understanding aprosodia per se. The main problem is that while subjects in laterality studies are selected for unilateral brain damage, they are not screened for prosodic impairment. Thus, the data pool on which we rely for conclusions about the nature of RHD prosodic deficits consists largely of subjects with and subjects without prosodic impairment. The characteristics of aprosodia, its mechanisms, duration, frequency of occurrence in the general RHD population, and the presence/absence of other RHD deficits that may accompany it have yet to be clearly delineated. These issues will remain unclear until a working definition of aprosodia is established and descriptive studies using that definition as a means of screening patients are undertaken.

See also PROSODIC DEFICITS, RIGHT HEMISPHERE LANGUAGE AND COMMUNICATION FUNCTIONS IN ADULTS; RIGHT HEMISPHERE LANGUAGE DISORDER.

—Penelope S. Myers

References


**Further Reading**


**Augmentative and Alternative Communication Approaches in Adults**

An augmentative and alternative communication (AAC) system is an integrated group of components used by individuals with severe communication disorders to enhance their competent communication. Competent communication serves a variety of functions, of which we can isolate four: (1) communication of wants and needs, (2) information transfer, (3) social closeness, and (4) social etiquette (Light, 1988). These four functions broadly encompass all communicative interactions. An appropriate AAC system addresses not only basic communication of wants and needs, but also the establishment, maintenance, and development of interpersonal relationships using information transfer, social closeness, and social etiquette.

AAC is considered multimodal, and as such it incorporates the full communication abilities of the adult. It includes any existing natural speech or vocalizations, gestures, formal sign language, and aided communication. “AAC allows individuals to use every mode possible to communicate” (Light and Binger, 1998, p. 1).

AAC systems are typically described as high-technology, low- or light-technology, and no-technology in respect to the aids used in implementation. High-technology AAC systems use electronic devices to support digitized or synthesized communication strategies. Low- or light-technology systems include items such as communication boards (symbols), communication books, and light pointing devices. A no-technology system involves the use of strategies and techniques, such as body movements, gestures, and sign language, without the use of specific aids or devices.

AAC is used to assist adults with a wide range of disabilities, including congenital disabilities (e.g., cerebral palsy, mental retardation), acquired disabilities (e.g., traumatic head injury, stroke), and degenerative conditions (e.g., multiple sclerosis, amyotrophic lateral sclerosis) (American Speech-Language-Hearing Association [ASHA], 1989). Individuals at any point across the life span and in any stage of communication ability may use AAC (see the companion entry, AUGMENTATIVE AND ALTERNATIVE COMMUNICATION APPROACHES IN CHILDREN).

Adults with severe communication disorders benefit from AAC. ASHA (1991, p. 10) describes these people as “those for whom natural gestural, speech, and/or written communication is temporarily or permanently
inadequate to meet all of their communication needs.’’ An important consideration is that “although some individuals may be able to produce a limited amount of speech, it is inadequate to meet their varied communication needs” (ASHA, 1991, p. 10). AAC may also be used to support comprehension and cognitive abilities by capitalizing on residual skills and thus facilitating communication.

Many adults with severe communication disorders demonstrate some ability to communicate using natural speech. Natural speech is more time-efficient and linguistically flexible than other modes (involving AAC). Speech supplementation AAC techniques (alphabet and topic supplementation) used in conjunction with natural speech can provide extensive contextual knowledge to increase the listener’s ability to understand a message. As the quality of the acoustic signal and the quality of environmental information improve, comprehensibility—intelligibility in context—of messages is enhanced (Lindblom, 1990). Similarly, poor-quality acoustic signals and poor environmental information result in a deterioration in message comprehensibility. When a speaker experiences reduced acoustic speech quality, optimizing any available contextual information through AAC techniques will increase the comprehensibility of the message. “Given that communication effectiveness varies across social situations and listeners, it is important that individuals who use natural speech, speech-supplementation, and AAC strategies learn to switch communication modes depending on the situation and the listener” (Hustad and Beukelman, 2000, p. 103).

The patterns of communication disorders in adults vary from condition to condition. Persons with aphasia, traumatic brain injury, Parkinson’s disease, Guillain-Barré syndrome, multiple sclerosis, and numerous motor speech impairments benefit from using AAC (Beukelman and Mirenda, 1998). ACC approaches for a few adult severe communication disorders are described here.

Amyotrophic lateral sclerosis (ALS) is a disease of rapid degeneration involving the motor neurons of the brain and spinal cord that leaves cognitive abilities generally intact. The cause is unknown, and there is no known cure. For those whose initial impairments are in the brainstem, speech symptoms typically occur early in the disease progression. On average, speech intelligibility in this (bulbar) group declines precipitously approximately 10 months after diagnosis. For those whose impairment begins in the lower spine, speech intelligibility declines precipitously approximately 25 months after diagnosis. Some individuals maintain functional speech much longer. Clinically, a drop in speaking rate predicts the onset of the abrupt drop in speech intelligibility (Ball, Beukelman, and Pattee, 2001). As a group, 80% of individuals with ALS eventually require use of AAC. Because the drop in intelligibility is so sudden, intelligibility is not a good measure to use in determining the timing of an AAC evaluation. Rather, because the speaking rate declines more gradually, an AAC evaluation should be completed when an individual reaches 50% of his or her habitual speaking rate (approximately 100 words per minute) on a standard intelligibility assessment (such as the Sentence Intelligibility Test; Yorkston, Beukelman, and Tice, 1996). Frequent objective measurement of speaking rate is important to provide timely AAC intervention. Access to a communication system is increasingly important as ALS advances (Mathy, Yorkston, and Gutmann, 2000).

Traumatic brain injury (TBI) refers to injuries to the brain that involve rapid acceleration and deceleration, whereby the brain is whipped back and forth in a quick motion, which results in compromised neurological function (Levin, Benton, and Grossman, 1982). The goal of AAC in TBI is to provide a series of communication systems and strategies so that individuals can communicate at the level at which they are currently functioning (Doyle et al., 2000). Generally, recovery of cognitive functioning is categorized into phases (Blackstone, 1989). In the early phase, the person is minimally responsive to external stimuli. AAC goals include providing support to respond to one-step motor commands and discriminate one of an array of choices (objects, people, locations). AAC applications during this phase include low-technology pictures and communication boards and choices of real objects to support communication. In the middle phase, the person exhibits improved consistency of responses to external stimuli. It is in this phase that persons who are unable to speak because of severe cognitive confusion become able to speak. If they do not become speakers by the end of this phase, it is likely a result of chronic motor control and language impairments. AAC goals during this phase address providing a way to indicate basic needs and giving a response modality that increases participation in the evaluation and treatment process. AAC intervention strategies usually involve nonelectronic, low-technology, or no-technology interventions to express needs. In the late phase, if the person continues to be nonspeaking, it is likely the result of specific motor or language impairment. AAC intervention may be complicated by coexisting cognitive deficits. Intervention goals address provision of functional ways to interact with listeners in a variety of settings and to assist the individual to participate in social, vocational, educational, and recreational settings. AAC intervention makes use of both low- and high-technology strategies in this phase of recovery.

Brainstem stroke (cerebrovascular accident, or CVA) disrupts the circulation serving the lower brainstem. The result is often severe dysarthria or anarthria, and reduced ability to control the muscles of the face, mouth, and larynx voluntarily or reflexively. Communication symptoms vary considerably with the extent of damage. Some individuals are dysarthric but able to communicate partial or complete messages, while others may be unable to speak. AAC intervention is typically described in five stages (Beukelman and Mirenda, 1998). In stage 1, the person exhibits no functional speech. The goal of intervention is to provide early communication so that
the person can respond to yes/no questions, initial choice making, pointing, and introduction of a multipurpose AAC device. In stage 2, the goal is to reestablish speech by working directly to develop control over the respiratory, phonatory, velopharyngeal, and articulatory sub-systems. Early in this stage, the AAC system will support the majority of interactions; however, late in this stage persons are able to convey an increasing percentage of messages with natural speech. In stage 3, the person exhibits independent use of natural speech. The AAC intervention focuses on intelligibility, with alphabet supplementation used early, but later only to resolve communication breakdowns. In this stage, the use of AAC may become necessary only to support writing. In stages 4 and 5, the person no longer needs to use an AAC system.

In summary, adults with severe communication disorders are able to take advantage of increased communication through the use of AAC. The staging of AAC interventions is influenced by the individual’s communication abilities and the natural course of the disorder, whether advancing, remitting, or stable.

—Laura J. Ball

References


Further Readings


Augmentative and Alternative Communication Approaches in Children

The acquisition of communication skills is a dynamic, bidirectional process of interactions between speaker and listener. Children who are unable to meet their daily needs using their own speech require alternative systems to support their communication interaction efforts (Reichle, Beukelman, and Light, 2001). An augmentative and alternative communication (AAC) system is an
integrated group of components used by a child to enhance or develop competent communication. It includes any existing natural speech or vocalizations, gestures, formal sign language, and aided communication. “AAC allows individuals to use every mode possible to communicate” (Light and Drager, 1998, p. 1).

The goal of AAC support is to provide children with access to the power of communication, language, and literacy. This power allows them to express their needs and wants, develop social closeness, exchange information, and participate in social, educational, and community activities (Beukelman and Mirenda, 1998). In addition, it provides a foundation for language development and facilitates literacy development (Light and Drager, 2001). Timeliness in implementing an AAC system is paramount (Reichle, Beukelman, and Light, 2001). The earlier that graphic and gestural mode supports can be put into place, the greater will be the child’s ability to advance in communication development.

Children experience significant cognitive, linguistic, and physical growth throughout their formative years, from preschool through high school. AAC support for children must address both their current communication needs as well as predict future communication needs and abilities, so that they will be prepared to communicate effectively as they mature. Because participation in the general classroom requires many kinds of extensive communication, effective AAC systems that are age appropriate and context appropriate serve as critical tools for academic success (Sturm, 1998, p. 391). Early interventions allow children to develop the linguistic, operational, and social competencies necessary to support their participation in academic settings.

Many young children and those with severe multiple disabilities cannot use traditional spelling and reading skills to access their AAC systems. Very young children, who are preliterate, have not yet developed reading and writing skills, while older children with severe cognitive impairments may remain nonliterate. For individuals who are not literate, messages within their AAC systems must be represented by one or more symbols or codes. With children, early communication development focuses on vocabulary that is needed to communicate essential messages and to develop language skills. Careful analysis of environmental and communication needs is used to develop vocabulary for the child’s AAC system. This vocabulary selection assessment includes examination of the ongoing process of vocabulary and message maintenance.

The vocabulary needs of children comprise contextual variations, including school talk, in which they speak with relatively unfamiliar adults in order to acquire knowledge. Home talk is used with familiar persons to meet needs and develop social closeness, as well as to assist parents in understanding their child. An example of vocabulary needs is exhibited by preschool children, who have been found to use generic small talk for nearly half of their utterances, when in preschool and at home (Ball et al., 1999). Generic small talk refers to messages that can be used without change in interaction with a variety of different listeners. Examples include “Hello”; “What are you doing?”; “What's that?”; “I like that!”; and “Leave me alone!”

Extensive instructional resources are available to school-age children. In the United States, the federal government has mandated publicly funded education for children with disabilities, in the form of the Individuals with Disabilities Education Act, and provides a legal basis for AAC interventions. Public policy changes have been adopted in numerous other countries to address resources available to children with disabilities.

AAC interventionists facilitate transitions from the preschool setting to the school setting by ensuring comprehensive communication through systematic planning and establishing a foundation for communication. A framework for integrating children into general education programs may be implemented by following the participation model (Beukelman and Mirenda, 1998), which includes four variables that can be manipulated to achieve appropriate participation for any child. Children transitioning from preschool to elementary school, self-contained to departmentalized programs, or school to post-school (vocational) will attain optimal participation when consideration is made for integration, social participation, academic participation, and independence. An AAC system must be designed to support literacy and other academic skill development as well as peer interactions. It must be appealing to children so that they find the system attractive and will continue using it (Light and Drager, 2001).

AAC systems are used by children with a variety of severe communication disorders. Cerebral palsy is a developmental neuromuscular disorder resulting from a nonprogressive abnormality of the brain. Children with severe cerebral palsy primarily experience motor control problems that impair their control of their speech mechanisms. The resulting motor speech disorder (dysarthria) may be so severe that AAC technology is required to support communication. Large numbers of persons with cerebral palsy successfully use AAC technology (Beukelman, Yorkston, and Smith, 1985; Mirenda and Mathy-Laikko, 1989). Typically, AAC support is provided to these children by a team of interventionists. In addition to the communication/AAC specialist, the primary team often includes occupational and physical therapists, technologists, teachers, and parents. A secondary support team might include orthotists, rehabilitation engineers, and pediatric ophthalmologists.

Intellectual disability, or mental retardation, is characterized by significantly subaverage intellectual functioning coexisting with limitations in adaptive skills (communication, self-care, home living, social skills, community use, self-direction, health and safety, academics, leisure, and work) that appear before the age of 18 (Luckasson et al., 1992). For children with communication impairments, it is important to engage in AAC instruction and interactions in natural rather than segregated environments. Calculator and Bedrosian noted that “communication is neither any more nor less than
a tool that facilitates individuals’ abilities to function in the various activities of daily living” (1988, p. 104). Children who are unable to speak because of cognitive limitations, with and without accompanying physical impairments, have demonstrated considerable success using AAC strategies involving high-technology (electronic devices) and low-technology (communication boards and books) options (Light, Collier, and Parnes, 1985a, 1985b, 1985c).

Autism and pervasive developmental disorders are described with three main diagnostic features: (1) impaired social interaction, (2) impaired communication, and (3) restricted, repetitive, and stereotypical patterns of behaviors, interests, and activities (American Psychiatric Association, 1994). These disorders occur as a spectrum of impairments of different causes (Wing, 1996). Children with a pervasive developmental disorder may have cognitive, social/communicative, language, and processing impairments. Early intervention with an emphasis on speech, language, and communication is extremely important (Dawson and Osterling, 1997). A range of intervention approaches has been suggested, and as a result, AAC interventionists may need to work with professionals whose views differ from their own, thus necessitating considerable collaboration (Simonsen, Olley, and Rosenthal, 1987; Dawson and Osterling, 1997; Freeman, 1997).

Developmental apraxia of speech (DAS) results in language delays, communication problems that influence academic performance, communication problems that limit effective social interaction, and significant speech production disorder. Children with suspected DAS have difficulty performing purposeful voluntary movements for speech (Caruso and Strand, 1999). Their phonological systems are impaired because of their difficulties in managing the intense motor demands of connected speech (Strand and McCauley, 1999). Children with DAS have a guarded prognosis for the acquisition of intelligible speech (Bernthal and Bankson, 1993). DAS-related speech disorders may result in prolonged periods of unintelligibility, particularly during the early elementary grades.

There is ongoing debate over the best way to manage suspected DAS. Some children with DAS have been treated with phonologically based interventions and others with motor learning tasks. Some interventionists support very intense schedules of interventions. These arguments have changed little in the last 20 years. However, the need to provide these children with some means to communicate so that they can successfully participate socially and in educational activities is becoming increasingly accepted. Cumley (1997) studied children with DAS who were provided with AAC technology. He reported that the group of children with lower speech intelligibility scores used their AAC technology more frequently than children with higher intelligibility. When children with DAS with low intelligibility scores used AAC technology, they did not reduce their speech efforts, but rather used the technology to resolve communication breakdowns. The negative effect of reduced speech intelligibility on social and educational participa-

tion has been documented extensively (Kent, 1993; Camarata, 1996). The use of AAC strategies to support the communicative interactions of children with such severe DAS that their speech is unintelligible is receiving increased attention (Culp, 1989; Cumley and Swanson, 2000).

In summary, children with severe communication disorders benefit from using AAC systems, from a variety of perspectives. Children with an assortment of clinical disorders are able to take advantage of increased communication through the use of AAC. The provision of AAC intervention is influenced by the child’s communication abilities and access to memberships. Membership involves integration, social participation, academic participation, and ultimately independence. A web site (http://aac.unl.edu) provides current information about AAC resources for children and adults.

—Laura J. Ball

References


Further Readings


Autism

The term *autism* was first used in 1943 by Leo Kanner to describe a syndrome of “disturbances in affective contact,” which he observed in 11 boys who lacked the dysmorphology often seen in mental retardation, but who were missing the social motivation toward communication and interaction that is typically present even in children with severe intellectual deficits. Despite their obvious impairments in social communication, the children Kanner observed did surprisingly well on some parts of IQ tests, leading Kanner to believe they did not have mental retardation.

Kanner’s observation about intelligence has been modified by subsequent research. When developmentally appropriate, individually administered IQ testing is administered, approximately 80% of people with autism score in the mentally retarded range, and scores remain stable over time (Rutter et al., 1994). However, individuals with autism do show unusual scatter in their abilities, with nonverbal, visually based performance often significantly exceeding verbal skills; unlike the performance seen in children with other kinds of retardation, whose scores are comparable across all kinds of tasks.
Recent research on the genetics of autism suggests that there are heritable factors that may convey susceptibility (Rutter et al., 1997). This vulnerability may be expressed in a range of social, communicative, and cognitive difficulties expressed in varying degrees in parents, siblings, and other relatives of individuals with autism.

Although genetic factors appear to contribute to some degree to the appearance of autism, the condition can also be associated with other medical conditions. Dykens and Volkmar (1997) reported the following:

- Approximately 25% of individuals with autism develop seizures.
- Tuberous sclerosis (a disease characterized by abnormal tissue growth) is associated with autism with higher than expected prevalence.
- The co-occurrence of autism and fragile X syndrome (the most common heritable form of mental retardation) is also higher than would be expected by chance.

Autism is considered one of a class of disabilities referred to as pervasive developmental disorders, according to the Diagnostic and Statistical Manual of the American Psychiatric Association (4th ed., 1994). The diagnostic criteria for autism are more explicitly stated in DSM-IV than the criteria for other pervasive developmental disorders. The criteria for autism are the result of a large field study conducted by Volkmar et al. (1994). The field trial showed that the criteria specified in DSM-IV exhibit reliability and temporal stability. Similar research on diagnostic criteria for other pervasive developmental disorders is not yet available.

The primary diagnostic criteria for autism include the following:

**Early onset.** Many parents first become concerned at the end of the first year of life, when a child does not start talking. At this period of development, children with autism also show reduced interest in other people; less use of communicative gestures such as pointing, showing, and waving; and noncommunicative sound making, perhaps including echoing that is far in advance of what can be produced in spontaneous or meaningful contexts. There may also be unusual preoccupations with objects (e.g., an intense interest in vacuum cleaners) or actions (such as twanging rubber bands) that are not like the preoccupations of other children at this age.

**Impairment in social interaction.** This is the hallmark of the autistic syndrome. Children with autism do not use facial expressions, eye contact, body posture, or gestures to engage in social interaction as other children do. They are less interested in sharing attention to objects and to other people, and they rarely attempt to direct others’ attention to objects or events they want to point out. They show only fleeting interest in peers, and often appear content to be left on their own to pursue their solitary preferred activities.

**Impairment in communication.** Language and communicative difficulties are also core symptoms in autism. Communicative differences in autism include the following:

- Mutism. Approximately half of people with autism never develop speech. Nonverbal communication, too, is greatly restricted (Paul, 1987). The range of communicative intentions expressed is limited to requesting and protesting. Showing off, labeling, acknowledging, and establishing joint attention, seen in normal preverbal children, are absent in this population. Wants and needs are expressed preverbally, but forms for expression are aberrant. Some examples are pulling a person toward a desired object without making eye contact, instead of pointing, and the use of maladaptive and self-injurious behaviors to express desires (Donnellan et al., 1984). Pointing, showing, and turning-taking are significantly reduced.

- For people with autism who do develop speech, both verbal and nonverbal forms of communication are impaired. Forty percent of people with autism exhibit echolalia, an imitation of speech they have heard—either immediate echolalia, a direct parroting of speech directed to them, or delayed echolalia, in which they repeat snatches of language they have heard earlier, from other people or on TV, radio, and so on. Both kinds of echolalia are used to serve communicative functions, such as responding to questions they do not understand (Prizant and Duchan, 1981). Echolalia decreases, as in normal development, with increases in language comprehension.

A significant delay in comprehension is one of the strongest distinctions between people with autism and those with other developmental disabilities (Rutter, Maywood, and Howlin, 1992). Formal aspects of language production are on par with developmental level. Children with autism are similar to mental age–matched children in the acquisition of rule-governed syntax, but language development lags behind nonverbal mental age (Lord and Paul, 1997). Articulation is on par with mental age in children with autism who speak; however, high-functioning adults with autism show higher than expected rates of speech distortions (Shriberg et al., 2001).

Word use is a major area of deficit in those who speak (Tager-Flusberg, 1995). Words are assigned to the same categories that others use (Minshew and Goldstein, 1993), and scores on vocabulary tests are often a strength. However, words may be used with idiosyncratic meanings, and difficulty is seen with deictic terms (i.e., you/I, here/there), whose meaning changes, depending on the point of view of the speaker). This was first thought to reflect a lack of self, as evidenced by difficulty with saying I. More recent research suggests that the flexibility required to shift referents and difficulty assessing others’ state of knowledge are more likely to account for this observation (Lee, Hobson, and Chiat, 1994).

Pragmatic, interpersonal uses of language present the greatest challenges to speakers with autism. The rate of initiation of communication is low (Stone and Caro-Martinez, 1990), and speech is often idiosyncratic and contextually inappropriate (Lord and Paul, 1997). Few references are made to mental states, and
people with autism have difficulty inferring the mental states of others (Tager-Flusberg, 1995). Deficits are seen in providing relevant responses or adding new information to established topics; primitive strategies such as imitation are used to continue conversations (Tager-Flusberg and Anderson, 1991).

For individuals at the highest levels of functioning, conversation is often restricted to obsessive interests. There is little awareness of listeners’ lack of interest in extended talk about these topics. Difficulty is seen in adapting conversation to take into account all participants’ purposes. Very talkative people with autism are impaired in their ability to use language in functional, communicative ways (Lord and Paul, 1997), unlike other kinds of children with language impairments, whose language use improves with increased amount of speech.

Paralinguistic features such as voice quality, intonation, and stress are frequently impaired in speakers with autism. Monotonic intonation is one of the most frequently recognized aspects of speech in autism. It is a major contributor to listeners’ perception of oddness (Mesibov, 1992). The use of pragmatic stress in spontaneous speech and speech fluency are also impaired (Shriberg et al., 2001).

Stereotypic patterns of behavior. Abnormal preoccupations with objects or parts of objects are characteristic of autism, as is a need for routines and rituals always to be carried out in precisely the same way. Children with autism become exceedingly agitated over small changes in routine. Stereotyped motor behaviors, such as hand flapping, are also typical but are related to developmental level and are likely to emerge in the preschool period.

Delays in imaginative play. Children with autism are more impaired in symbolic play behaviors than in other aspects of cognition, although strengths are seen in constructive play, such as stacking and nesting (Schuler, Prizant, and Wetherby, 1997).

There is no medical or biological profile that can be used to diagnose autism, nor is there one diagnostic test that definitely identifies this syndrome. Current assessment methods make use primarily of multidimensional scales, either interview or observational, that provide separate documentation of aberrant behaviors in each of the three areas that are known to be characteristic of the syndrome: social reciprocity, communication, and restricted, repetitive behaviors. The most widely used for research purposes are the Autism Diagnostic Interview (Lord, Rutter, and Le Conteur, 1994) and the Autism Diagnostic Observation Scale (Lord et al., 2000).

Until recently, autism was thought to be a rare disorder, with prevalence estimates of 4–5 per 10,000 (Lotter, 1966). However, these prevalence figures were based on identifying the disorder in children who, like the classic patients described by Kanner, had IQs within normal range. As it became recognized that the social and communicative deficits characteristic of autism could be found in children along the full range of the IQ spectrum, prevalence estimates rose to 1 per 1000 (Bryson, 1997).

Currently, there is a great deal of debate about incidence and prevalence, particularly about whether incidence is rising significantly. Although clinicians see more children today who receive a label of autism than they did 10 years ago, this is likely to be due to a broadening of the definition of the disorder to include children who show some subset of symptoms without the full-blown syndrome. Using this broad definition, current prevalence estimates range from 1 in 500 to as low as 1 in 300 (Fombonne, 1999). Although there is some debate about the precise ratio, autism is more prevalent in males than in females (Bryson, 1997).

In the vast majority of cases, children with autism grow up to be adults with autism. Only 1%–2% of cases have a fully normal outcome (Paul, 1987). The classic image of the autistic child—mute or echolalic, with stereotypic behaviors and a great need to preserve sameness—is most characteristic of the preschool period. As children with autism grow older, they generally progress toward more, though still aberrant, social involvement. In adolescence, 10%–35% of children with autism show some degree of regression (Gillberg and Schaumann, 1981). Still, with continued intervention, growth in both language and cognitive skills can be seen (Howlin and Goode, 1998). Approximately 75% of adults with autism require high degrees of support in living, with only about 20% gainfully employed (Howlin and Goode, 1998). Outcome in adulthood is related to IQ, with good outcomes almost always associated with IQs above 60 (Rutter, Greenfield, and Lockyer, 1967). The development of functional speech by age 5 is also a strong predictor of good outcome (DeMyer et al., 1973).

Major changes have taken place in the treatments used to address autistic behaviors. Although a variety of pharmacological agents have been tried, and some are effective at treating certain symptoms (see McDougle, 1997, for a review), the primary forms of treatment for autism are behavioral and educational. Early intervention, when provided with a high degree of intensity (at least 20 hours per week), has proved particularly effective (Rogers, 1996). There is ongoing debate about the best methods of treatment, particularly for lower functioning children. There are proponents of operant applied behavior treatments (Lovas, 1987), of naturalistic child-centered approaches (Greenspan and Wieder, 1997), and of approaches that are some hybrid of the two (Prizant and Wetherby, 1998). Recent innovations focus on the use of alternative communication systems (e.g., Bondi and Frost, 1998) and on the use of environmental compensatory supports, such as visual calendars, to facilitate communication and learning (Quill, 1998). Although all of these approaches have been shown to be associated with growth in young children with autism, no definitive study has yet compared approaches or measured long-term change.

For higher functioning and older individuals with autism, most interventions are derived from more
general strategies used in children with language impairments. These strategies focus on the development of conversational skills, the use of scripts to support communication, strategies for communicative repair, and the use of reading to support social interaction (Prizant et al., 1997). “Social stories,” in which anecdotal narratives are encouraged to support social understanding and participation, is a new method that is often used with higher functioning individuals (Gray, 1995).

—Rhea Paul

References


Further Readings


---

Bilingualism, Speech Issues in

In evaluating the properties of bilingual speech, an anterior question that must be answered is who qualifies as bilingual. Scholars have struggled with this question for decades. Bilingualism defies delimitation and is open to a variety of descriptions and interpretations. For example, Bloomfield (1933) required native-like control of two languages, while Weinreich (1968) and Mackey (1970) considered as bilingual an individual who alternately used two languages. Beatens-Beardsmore (1982), observing a wide range of variations in different contexts, concluded that it is not possible to formulate a single neat definition, and stated that bilingualism as a concept has “open-ended semantics.”

It has long been recognized that bilingual individuals form a heterogeneous population in that their abilities in their two languages are not uniform. Although some bilingual speakers may have attained a native-like production in each language, the great majority are not balanced between the two languages. The result is interference from the dominant language. Whether a child becomes bilingual simultaneously (two languages are acquired simultaneously) or successively (one language, generally the home language, is acquired earlier, and the other language is acquired, for example, when the child goes to school), it is impossible to rule out interference. In the former case, this may happen because of different degrees of exposure to the two languages; in the latter, the earlier acquired language may put its imprint on the one acquired later. More children become bilingual successively, and the influence of one language on the other is more evident. Yet even here there is no uniformity among speakers, and the range of interference from the dominant language forms a continuum.

The different patterns that a bilingual child reveals in speech may not necessarily be the result of interference. Bilingual children, like their monolingual counterparts, may suffer speech and/or language disorders. Thus, when children who grow up with more than one language produce patterns that are erroneous with respect to the speech of monolingual speakers, it is crucial to determine whether these nonconforming patterns are due to the influence of the child’s other language or are indications of a speech-language disorder.

To be able to make accurate diagnoses, speech-language pathologists must use information from interference patterns, normal and disordered phonological development in general as well as in the two languages, and the specific dialect features. In assessing the phonological development of a bilingual child, both languages should be the focus of attention, and each should be examined in detail, even if the child seems to be a dominant speaker of one of the languages. To this end, all phonemes of the languages should be assessed in different word positions, and phonotactic patterns should be evaluated. Assessment tools that are designed for English, no matter how perfect they are, will not be
appropriate for the other language and may be the cause of over- or underdiagnosis.

Certain cases lend themselves to obvious identification of interference. For example, if we encounter in the English language productions of a Portuguese-English bilingual child forms like [t iz] for “tease” and [t ip] for “tip,” whereby target /t/ turns into [t f], we can, with confidence, say that these renditions were due to Portuguese interference, as such substitutions are not commonly observed in developmental phonologies, and the change of /t/ to [t f] before /i/ is a rule of Portuguese phonology.

The decision is not always so straightforward, however. For example, substitutions may reflect certain developmental simplification processes that are universally phonetically motivated and shared by many languages. If a bilingual child’s speech reveals any such processes, and if the first language of the child does not have the opportunities for such processes to surface, then it would be very difficult to identify the dominant language as the culprit and label the situation as one of interference. For example, if a 6-year-old child bilingual in Spanish and English reveals processes such as final obstruent devoicing (e.g., [bæk] for “bag,” [bet] for “bed”) and/or deletion of clusters that do not follow sonority sequencing ([tap] for “stop,” [pot] for “spit”), we cannot claim that these changes are due to Spanish interference. Rather, these processes are among the commonly occurring developmental processes that occur in the speech of children in many languages. However, because these common simplification processes are usually suppressed in normally developing children by age 6, this particular situation suggests a delay or disorder. In this case, these processes may not have surfaced until age 6 because none of these patterns are demanded by the structure of Spanish. In other words, because Spanish has no voiced obstruents in final position and no consonant clusters that do not follow sonority sequencing, it is impossible to refer to the first language as the explanation. In such instances we must attribute these patterns to universally motivated developmental processes that have not been eliminated according to the expected timetable.

We may also encounter a third situation in which the seemingly clear distinction between interference and the developmental processes is blurred. This occurs when one or more of the developmental processes are also the patterns followed by the first (dominant) language. An example is final obstruent devoicing in the English language productions of a child with German, Russian, Polish, or Turkish as the first language. Although final obstruent devoicing is a natural process that even occurs in the early speech of monolingual English-speaking children, it is also a feature of the languages listed. Thus, the result is a natural tendency that receives extra impetus from the rule of the primary system. Other examples that could be included in the same category would be consonant cluster reduction in children whose primary language is Japanese, Turkish, or Finnish, and single obstructive coda deletion in children whose primary language is Japanese, Italian, Spanish, or Portuguese.

Besides the interference patterns and common developmental processes, speech-language pathologists must be watchful for some unusual (idiosyncratic) processes that are observed in children (Grunwell, 1987; Dodd, 1993). Processes such as unusual cluster reduction, as in [ren] for train (instead of the expected [ten]), fricative gliding, as in [wəg] for fig, frication of stops, as in [vən] for ban, and backing, as in [pæk] for pat, may occur in children with phonological disorders.

Studies that have examined the phonological patterns in normally developing bilingual children (Gildersleeve, Davis, and Stubbe, 1996) and bilingual children with a suspected speech disorder (Dodd, Holm, and Wei, 1997) indicate that children in both groups exhibit patterns different from matched, monolingual peers. Compared with their monolingual peers, normally developing bilingual children and bilingual children with phonological disorders had a lower overall intelligibility rating, made more errors overall, distorted more sounds, and produced more uncommon error patterns. As for the difference between normally developing bilingual children and bilingual children with phonological disorders, it appears that children with phonological disorders manifest more common simplification patterns, suppress such patterns over time more slowly, and are likely to have uncommon processes.

As speech-language pathologists become more adept at differentiating common and uncommon phonological patterns and interference patterns in bilingual children, they will also need to consider not only the languages of the client, but also the specific dialects of those languages. Just as there are several varieties of English spoken in different countries (e.g., British, American, Australian, South African, Canadian, Indian) and even within one country (New England variety, Southern variety, General American, and African American Vernacular in the United States), other languages also show dialectal variation. Because none of these varieties or dialects of a given language is or can be considered a disordered form of that language, the child’s dialectal information is essential. Any assessment of the child’s speech must be made with respect to the norm of the particular variety she or he is learning. Not accounting for dialect features may either result in the misdiagnosis of a phonological disorder or escalate the child’s severity rating.

Last but definitely not least is the desperate need for information on phonological development in bilingual children and assessment procedures unique to these individuals. Language skills in bilingual persons have almost always been appraised in reference to monolingual standards (Grosjean, 1992). Accordingly, a bilingual child is assessed with two procedures, one for each language, that are designed to evaluate monolingual speakers of these languages. This assumes that a bilingual individual is two monolingual individuals in one person. However, because of the constant interaction of the two languages, each phonological system of a bilin-
gual child may, and in most cases will, not necessarily be acquired in a way identical to that of a monolingual child (Watson, 1991).

In order to characterize bilingual phonology accurately, detailed information on both languages being acquired by the children is indispensable. However, data on the developmental patterns in two languages separately would not be adequate, as information on phonological development in bilingual children is the real key to understanding bilingual phonology. Because bilingual speakers’ abilities in the two languages vary immensely from one individual to another, developing assessment tools for phonological development is a huge task, perhaps the biggest challenge for the field.

— Mehmet Yavas

References


Further Readings


So, L. (1992). Cantonese Segmental Phonology Test. Hong Kong University, Department of Speech and Hearing Sciences.


Robin, 1993), phonologically based deficits in representation (e.g., Velleman and Strand, 1993), or deficits in neural tissue with organizational consequences (e.g., Crary, 1984; Sussman, 1988) have been posited. Reflecting these varied views of causality, a variety of terms have been employed: developmental apraxia of speech, developmental verbal dyspraxia, and developmental articulatory dyspraxia. Clinically, DAS has most often been defined by exclusion from functional speech disorder or delay using a complex of behavioral symptoms (e.g., Stackhouse, 1992; Shriberg, Aram, and Kwiatkowski, 1997).

The characterization of DAS was originally derived from apraxia of speech in adults, a disorder category based on acquired brain damage resulting in difficulty in programming speech movements (Broca, 1861). Morley, Court, and Miller (1954) first applied the term dyspraxia to children based on a proposed similarity in behavioral correlates with adult apraxic symptoms. A neurological etiology was implied by the analogy but has not been conclusively delineated, even with increasingly sophisticated instrumental techniques for understanding brain-behavior relations (see Bennett and Netsell, 1999; LeNormand et al., 2000). Little coherence and consensus is available in the literature at present. In addition, despite nearly 40 years of research, differential diagnostic correlates and range of severity levels characterizing DAS remain imprecisely defined. Guyette and Deidrich (1981) have suggested that DAS may not be a theoretically or clinically definable entity, as current empirical evidence does not produce any behavioral symptom not overlapping with other categories of developmental speech disorder or delay. In addition, no currently available theoretical constructs specifically disprove other possible theories for the origins of DAS (see Davis, Jakielski, and Marquardt, 1998). In contrast to developmental disorder categories such as hearing impairment or cleft palate, lack of a link of underlying cause or theoretical base with behavioral correlates results in an “etiological” disorder label with no clearly established basis. Evidence for a neurological etiology for DAS is based on behavioral correlates that are ascribed to a neurological basis, thus achieving a circular argument structure for neural origins (Marquardt, Sussman, and Davis, 2000).

Despite the lack of consensus on theoretical motivation, etiology, or empirical evidence precisely defining behavioral correlates, there is some consensus among practicing clinicians as well as researchers (e.g., Shriberg, Aram, and Kwiatkowski, 1997) that DAS exists. It thus represents an incompletely understood disorder that poses important challenges both to practicing clinicians and to the establishment of a consistent research base for overall understanding. An ethical differential diagnosis for clinical intervention and research investigations should, accordingly, be based on awareness of the current state of empirically established data regarding theories and behavioral correlates defining this disorder. Cautious application of the diagnostic label should be the norm, founded on a clear understanding of positive benefits to the client in discerning long-term prognosis, appropriate decisions regarding clinical intervention, and valid theory building to understand the underlying nature of the disorder. Use of DAS as an “umbrella term for children with persisting and serious speech difficulties in the absence of obvious causation, regardless of the precise nature of their unintelligibility” (Stackhouse, 1992, p. 30) is to be avoided. Such practice continues to cloud the issue of precise definition of the presence and prevalence of the disorder in child clinical populations.

Accordingly, a review of the range of behavioral correlates presently in use is of crucial importance to careful definition and understanding of DAS. The relationship of behavioral correlates to differential diagnosis from “functional” speech disorder or delay is of primary importance to discriminating DAS as a subcategory of functional speech disorder. If no single defining characteristic or complex of characteristics emerges to define DAS, the utility of the label is seriously questionable for either clinical or research purposes. In every instance, observed behaviors need to be evaluated against developmental behaviors appropriate to the client’s chronological age. In the case of very young clients, the differential diagnosis of DAS is complicated (Davis and Velleman, 2000). Some listed characteristics may be normal aspects of earliest periods of speech and language development (e.g., predominant use of simple syllable shapes or variability in production patterns at the onset of meaningful speech; see Vihman, 1997, for a review of normal phonetic and phonological development).

Before the clinical symptoms presently employed to define DAS are outlined, specific issues with available research will be reviewed briefly. It should be emphasized that behavioral inclusion criteria are not consistently reported and differing criteria are included across studies. Criteria for inclusion in studies then become recognized symptoms of involvement, achieving a circularity that is not helpful for producing valid characterization of the disorder (Stackhouse, 1992). Subject ages vary widely, from preschoolers (Bradford and Dodd, 1996) to adults (Ferry, Hall, and Hicks, 1975). Some studies include control populations of functional speech disorders for differential diagnosis (Stackhouse, 1992; Dodd, 1995); others do not (Horowitz, 1984). Associated language and praxic behaviors are included as differential diagnostic correlates in some studies (Crary and Towne, 1984), while others explicitly exclude these deficits (e.g., Hall, Jordon, and Robin, 1993). Severity is not reported consistently. When it is reported, the basis for assigning severity judgments is inconsistent across studies. A consequence of this inconsistency is lack of consensus on severity level appropriate to the DAS label. In some reports, the defining characteristic is severe and persistent disorder (e.g., Shriberg, Aram, and Kwiatkowski, 1997). In other reports (e.g., Thoonen et al., 1997), a continuum of severity is explored. In the latter conceptualization, DAS can manifest as mild, moderate, or severe speech disorder.

Despite the foregoing critique, the large available literature on DAS suggests some consensus on behav-
ioral correlates that should be evaluated in establishing a differential diagnosis. The range of expression of these characteristics, although frequently cited, has not been specified quantitatively. Accordingly, these behaviors should not be considered definitive but suggestive of directions for future research as well as guidelines for the practicing clinician based on emerging research. Exclusionary criteria for a differential diagnosis have been suggested in the areas of peripheral motor and sensory function, cognition, and receptive language. Exclusionary criteria frequently noted include (1) no peripheral organic disorder (e.g., cleft palate), (2) no sensory deficit (i.e., in vision or hearing), (3) no peripheral muscle weakness or dysfunction (e.g., dysarthria, cerebral palsy), (4) normal IQ, and (5) normal receptive language.

Phonological and phonetic correlates have also been listed. Descriptive terminology varies from phonetic (e.g., Murdoch et al., 1995) to phonological (Forrest and Morrisette, 1999; Velleman and Shriberg, 1999) according to the theoretical perspective of the researcher, complicating understanding of the nature of the disorder and comparison across studies. In addition, behavioral correlates have been established across studies with highly varied subject pools and differing exclusionary criteria. The range of expression of symptoms is not established (i.e., what types and severity of suprasegmental errors are necessary or sufficient for the diagnosis?). Some characteristics are in common with functional disorders and thus do not constitute a differential diagnostic characteristic (i.e., how limited does the consonant or vowel repertoire have to be to express DAS?). In addition, not all symptoms are consistently reported as being necessary to a diagnosis of DAS (e.g., not all clients show “groping postures of the articulators”). Long-term persistence of clinical symptoms in spite of intensive therapy has also frequently been associated with DAS. Phonological/phonetic correlates reported include (1) limited consonant and vowel phonetic inventory, (2) predominant use of simple syllable shapes, (3) frequent omission of errors, (4) a high incidence of vowel errors, (5) altered suprasegmental characteristics (including rate, pitch, loudness, and nasality), (6) variability and lack of consistent patterning in speech output, (7) increased errors on longer sequences, (8) groping postures, and (9) lack of willingness or ability to imitate a model.

Co-occurring characteristics of DAS in several related areas have also been mentioned frequently. However, co-occurrence may be optional for a differential diagnosis, because these characteristics have not been consistently tracked across available studies. Co-occurring characteristics frequently cited include (1) delays in gross and fine motor skills, (2) poor volitional oral nonverbal skills, (3) inconsistent diadokokinetic rates, (4) delay in syntactic development, and (5) reading and spelling delays.

Clearly, DAS is a problematic diagnostic category for both research and clinical practice. Although it has long been a focus of research and a subject of intense interest to clinicians, little consensus exists on definition, etiology, and characterization of behavioral or neural correlates. Circularity in the way in which etiology and behavioral correlates have been described and studied does not lend to precision in understanding DAS. Research utilizing consistent subject selection criteria is needed to begin to link understanding of DAS to ethical clinical practices in assessment and intervention and to elucidate the underlying causes of this disorder.

—Barbara L. Davis

References


**Further Readings**


**Dialect, Regional**

Dialects or language varieties are a result of systematic, internal linguistic changes that occur within a language. Unlike accents, in which linguistic changes occur mainly at the phonological level, dialects reflect structural changes in phonology, morphology, and syntax, as well as lexical or semantic changes. The degree of mutual intelligibility that a speaker’s language has with a des-
ignited standard linguistic system is often used to distinguish dialect from language. Mutual intelligibility means that speakers of one dialect can understand speakers of another dialect.

Although the construct of mutual intelligibility is frequently employed to differentiate dialect from language, there are counterexamples. On one hand, speakers may have the same language, but the dialects may not be mutually intelligible. For example, Chinese has a number of dialects, such as Cantonese and Mandarin, each spoken in different geographical regions. Although Cantonese and Mandarin speakers consider these dialects, the two lack mutual intelligibility since those who speak only Cantonese do not easily understand those who speak only Mandarin, and vice versa. On the other hand, speakers may produce different languages but have mutual intelligibility. For instance, Norwegian, Swedish, and Danish are thought of as different languages, yet speakers of these languages can easily understand one another.

A dialect continuum or dialect continua may account for lack of mutual intelligibility in a large territory. A dialect continuum refers to a distribution of sequentially arranged dialects that progressively change speech or linguistic forms across a broad geographical area. Some speech shifts may be subtle, others may be more dramatic. Assume widely dispersed territories are labeled towns A, B, C, D, and E and are serially adjacent to one another, thereby creating a continuum. B is adjacent to A and C, C is adjacent to B and D, and so on. There will be mutual intelligibility between dialects spoken in A and B, between B and C, between C and D, and between D and E. However, the dialects of the two towns at the extremes, A and E, may not be mutually intelligible, owing to the continuous speech and language shifts that have occurred across the region. It is also possible that some of the intermediate dialects, such as B and D, may not be mutually intelligible.

Because different conditions influence dialects, it is not easy to discriminate dialect precisely from language. Using mutual intelligibility as a primary marker of distinction should be considered relative to the territories of interest. For example, in the United States, the concept of mutual intelligibility appears valid, whereas it is not completely valid in many other countries.

Dialects exist in all languages and are often discussed in terms of social or regional varieties. Social dialects represent a speaker’s social stratification within a given society or cultural group. Regional dialects are associated with geographical location or where speakers live. Regional and social dialects may co-occur within language patterns of the same speaker. In other words, social and regional dialects are not mutually exclusive.

Regional dialects constitute a unique cluster of language characteristics that are distributed across a specified geographical area. Exploration of regional dialectal systems is referred to as dialectology, dialect geography, or linguistic geography. For many years, dialects spoken in cities were thought of as prestigious. Therefore, in traditional dialect studies, data were mainly collected in rural areas. Surveys, questionnaires, and interview techniques were used as primary mechanisms of data collection. A field worker would visit an area and talk to residents using predetermined elicitation techniques that would encourage the speaker to produce the distinctive items of interest. The field worker would then manually note whether the individual’s speech contained the distinctive linguistic features of interest. These methods generated a number of linguistic atlases that contained linguistic maps displaying geographical distributions of language characteristics.

Data were used to determine where a selected set of features was produced and where people stopped using this same set of features. The selected features could include vocabulary, specific sounds, or grammatical forms. Lines, or isoglosses, were drawn on a map to indicate the existence of specific features. When multiple isoglosses, or a bundle of isoglosses, surround a specific region, this is used to designate dialect boundaries. The bundle of isoglosses on a linguistic map would indicate that people on one side produced a number of lexical items and grammatical forms that were different from the speech of those who lived outside the boundary. Theoretically, the dialect was more distinctive, with a greater amount of bundling.

After the 1950s, audio and, eventually, video recordings were made of speakers in designated regions. Recordings allow a greater depth of analysis because they can be repeatedly replayed. Concurrent with these technological advances, there was increased interest in urban dialects, and investigators began to explore diversity of dialects within large cities, such as Boston, Detroit, New York, and London. Technological developments also led to more quantitative studies. Strong statistical analysis (dialectometry) has evolved since the 1970s and allows the investigator to explore large data sets with large numbers of contrasts.

Several factors contribute to the formation of regional dialects. Among these factors are settlement and migration patterns. For instance, regional English varieties began to appear in the United States as speakers immigrated from different parts of Britain. Speakers from the eastern region of England settled in New England, and those from Ulster settled in western New England and in Appalachia. Each contributed different variations to the region in which they settled.

Regional dialect formation may also result from the presence of natural boundaries such as mountains, rivers, and swamps. Because it was extremely difficult to traverse the Appalachian mountain range, inhabitants of the mountains were isolated and retained older English forms that contribute some of the unique characteristics of Appalachian English. For example, the morphological a-prefix in utterances such as “He come a-running” or “She was a-tellin’ the story” appears to be a retention from older forms of English that were prevalent in the seventeenth century.

Commerce and culture also play important roles in influencing regional dialects, as can be observed in the unique dialect of people in Baltimore, Maryland. Speakers of “Bawlamerese” live in “Merlin” (Maryland), whose state capitol is “Napolis” (Annapolis),
located next to “Warshnin” (Washington, D.C.), and refer to Bethlehem Steel as “Bethlum.” Because the “Bethlum” mill, located in Fells Point, has been a primary employer of many individuals, language has evolved to discuss employment. Many will say they work “down a point” or “down a mill,” where the boss will “har and far” (hire and fire) people. While most people working “down a point” live in “Dundock” (Dundalk), some may live as far away as “Norf Abnew” (North Avenue), “Habberdy Grace” (Harve de Grace), or even “Klumya” (Columbia).

Two other types of geolinguistic variables are often associated with regional dialects. One variable is a set of linguistic characteristics that are unique to a geographical area or that occur only in that area. For instance, unique to western Pennsylvania, speakers say “youse” (you singular), “yens” (you plural) and “yens boomers” (a group of people). The second variable is the frequency of occurrence of regional linguistic characteristics in a specific geographic area. For example, the expression “take ‘er easy” is known throughout the United States, but mainly used in central and western Pennsylvania.

—Adele Proctor

**Further Readings**


**Dysarthrias: Characteristics and Classification**

The dysarthrias are a group of neurological disorders that reflect disturbances in the strength, speed, range, tone, steadiness, timing, or accuracy of movements necessary for prosodically normal, efficient and intelligible speech. They result from central or peripheral nervous system conditions that adversely affect respiratory, phonatory, resonatory, or articulatory speech movements. They are often accompanied by nonspeech impairments (e.g., dysphagia, hemiplegia), but sometimes they are the only manifestation of neurological disease. Their course can be transient, improving, exacerbating–remitting, progressive, or stationary.

Endogenous or exogenous events as well as genetic influences can cause dysarthrias. Their neurological bases can be present congenitally or they can emerge acutely, subacutely, or insidiously at any time of life. They are associated with many neurological conditions, but vascular, traumatic, and degenerative diseases are their most common cause in most clinical settings; neoplastic, toxic-metabolic, infectious, and inflammatory causes are also possible.

Although incidence and prevalence are not precisely known, dysarthria often is present in a number of frequently occurring neurological diseases, and it probably represents a significant proportion of all acquired neurological communication disorders. For example, approximately one-third of people with traumatic brain injury may be dysarthric, with nearly double that prevalence during the acute phase (Sarno, Buonaguro, and Levita, 1986; Yorkston et al., 1999). Dysarthria probably occurs in 50%–90% of people with Parkinson’s disease, with increased prevalence as the disease progresses (Logemann et al., 1978; Mlcoch, 1992), and it can be among the most disabling symptoms of the disease in some cases (Dewey, 2000). Dysarthria emerges very frequently during the course of amyotrophic lateral sclerosis (ALS) and may be among the presenting symptoms and signs in over 20% (Rose, 1977; Gubbay et al., 1985). It occurs in 25% of patients with lacunar stroke (Arboix and Marti-Vilata, 1990). In a large tertiary care center, dysarthria was the primary communication disorder in 46% of individuals with any acquired neurological disease seen for speech-language pathology evaluation over a 4-year period (Duffy, 1995).

The clinical diagnosis is based primarily on auditory perceptual judgments of speech during conversation, sentence repetition, and reading, as well as performance on tasks such as vowel prolongation and alternating motion rates (AMRs; for example, repetition of “puh,” “tuh,” and “kuh” as rapidly and steadily as possible). Vowel prolongation permits judgments about respiratory support for speech as well as the quality, pitch, and duration of voice. AMRs permit judgments about the rate and rhythm of repetitive movements and are quite useful in distinguishing among certain dysarthria types (e.g., they are typically slow but regular in spastic dysarthria, but irregular in ataxic dysarthria). Visual and physical examination of the speech mechanism at rest and during nonspeech responses (e.g., observations of asymmetry, weakness, atrophy, fasciculations, adventitious movements, pathological oral reflexes) and information from instrumental measures (e.g., acoustic, endoscopic, videofluorographic) often provide confirmatory diagnostic evidence.

Dysarthria severity can be indexed in several ways, but quantitative measures usually focus on intelligibility and speaking rate. The most commonly used intelligibility measures are the Computerized Assessment of Intelligibility in Dysarthric Speakers (Yorkston, Beukelman, and Traynor, 1984) and the Sentence Intelligibility Test (Yorkston, Beukelman, and Tice, 1996), but other measures are available for clinical and research purposes (Enderby, 1983; Kent et al., 1989).

A wide variety of acoustic, physiological, and anatomical imaging methods are available for assessment. Some are easily used clinically, whereas others are primarily research tools. Studies using them have often
yields results consistent with predictions about pathophysiology from auditory-perceptual classification, but discrepancies that have been found make it clear that correspondence between perceptual attributes and physiology cannot be assumed (Duffy and Kent, 2001).

Methods that show promise or that already have refined what we understand about the anatomical and physiological underpinnings of the dysarthrias include acoustic, kinematic, and aerodynamic methods, electromyography, electroencephalography, radiography, tomography, computer tomography, magnetic resonance imaging, functional magnetic resonance imaging, positron emission tomography, single-photon emission tomography, and magnetoencephalography (McNeil, 1997; Kent et al., 2001).

The dysarthrias can be classified by time of onset, course, site of lesion, and etiology, but the most widely used classification system in use today is based on the auditory-perceptual method developed by Darley, Aronson, and Brown (1969a, 1969b, 1975). Often referred to as the Mayo Clinic system, the method identifies dysarthria types, with each type representing a perceptually distinguishable grouping of speech characteristics that presumably reflect underlying pathophysiology and focus of lesion. The following summarizes the major types, their primary distinguishing perceptual attributes, and their presumed underlying localization and distinguishing neurophysiological deficit.

**Flaccid dysarthria** is due to weakness in muscles supplied by cranial or spinal nerves that innervate respiratory, laryngeal, velopharyngeal, or articulatory structures. Its specific characteristics depend on which nerves are involved. Trigeminal, facial, or hypoglossal nerve lesions are associated with impaired articulation of phonemes that rely on jaw, face, or lingual movement. Vagus nerve lesions can lead to hypernasality or weak pressure consonant production when the pharyngeal branch is affected or to breathiness, hoarseness, diplopia, stridor, or short phrases when the laryngeal branches are involved. When spinal respiratory nerves are affected, reduced loudness, short phrases, and alterations in breath patterning for speech may be evident. In general, unilateral lesions and lesions of a single nerve produce relatively mild deficits, whereas bilateral lesions or multiple nerve involvement can have devastating effects on speech.

**Spastic dysarthria** is usually associated with bilateral lesions of upper motor neuron pathways that innervate relevant cranial and spinal nerves. Its distinguishing characteristics are attributed to spasticity, and they often include a strained-harsh voice quality, slow rate, slow but regular speech AMRs, and restricted pitch and loudness variability. All components of speech production are usually affected.

**Ataxic dysarthria** is associated with lesions of the cerebellum or cerebellar control circuits. Its distinguishing characteristics are attributed primarily to incoordination, and they are perceived most readily in articulation and prosody. Characteristics often include irregular articulatory breakdowns, irregular speech AMRs, inappropriate variations in pitch, loudness, and duration, and sometimes excess and equal stress across syllables.

**Hypokinetic dysarthria** is associated with basal ganglia control circuit pathology, and its features seem mostly related to rigidity and reduced range of motion. Parkinson’s disease is the prototypic disorder associated with hypokinetic dysarthria, but other conditions can also cause it. Its distinguishing characteristics include reduced loudness, breathy-tight dysphonia, monopitch and monoloudness, and imprecise and sometimes rapid, accelerating, or “blurred” articulation and AMRs. Dysfluency and palilalia also may be apparent.

**Hyperkinetic dysarthria** is also associated with basal ganglia control circuit pathology. Unlike hypokinetic dysarthria, its distinguishing characteristics are a product of involuntary movements that interfere with intended speech movements. Its manifestations vary across several causal movement disorders, which can range from relatively regular and slow (tremor, palatal-laryngeal myoclonus), to irregular but relatively sustained (dystonia), to relatively rapid and predictable or unpredictable (chorea, action myoclonus, tics). These movements may be a nearly constant presence, but sometimes they are worse during speech or activated only during speech. They may affect any one or all levels of speech production, and their effects on speech can be highly variable. Distinguishing characteristics usually reflect regular or unpredictable variability in phrasing, voice, articulation, or prosody.

**Unilateral upper motor neuron dysarthria** has an anatomical rather than pathophysiological label because it has received little systematic study. It most commonly results from stroke affecting upper motor neuron pathways. Because the damage is unilateral, severity usually is rarely worse than mild to moderate. Its characteristics often overlap with varying combinations of those associated with flaccid, spastic, or ataxic dysarthria (Duffy and Folger, 1996; Hartman and Abbs, 1992).

**Mixed dysarthrias** reflect combinations of two or more of the single dysarthria types. They occur more frequently than any single dysarthria type in many clinical settings. Some diseases are associated only with a specific mix; for example, flaccid-spastic dysarthria is the only mix expected in ALS. Other diseases, because the locus of lesions they cause is less predictable (e.g., multiple sclerosis, traumatic brain injury), may be associated with virtually any mix. The presence of mixed dysarthria is very uncommon or incompatible with some diseases (e.g., myasthenia gravis is associated only with flaccid dysarthria), so sometimes the presence of a mixed dysarthria can make a particular disease an unlikely cause or raise the possibility that more than a single disease is present.

Because of their potential to inform our understanding of the neural control of speech, and because their prevalence in frequently occurring neurological diseases is high and their functional effects are significant, dysarthrias draw considerable attention from clinicians and researchers. The directions of clinical and more basic...
research are broad, but many current efforts are aimed at the following: refining the differential diagnosis and indices of severity; delineating acoustic and physiological correlates of dysarthria types and intelligibility; more precisely establishing the relationships among perceptual dysarthria types, neural structures and circuitry, and acoustic and pathophysiological correlates; and developing more effective treatments for the underlying impairments and functional limitations imposed by them. Advances are likely to come from several disciplines (e.g., speech-language pathology, speech science, neurology) working in concert to integrate clinical, anatomical, and physiological observations into a coherent understanding of the clinical disorders and their underpinnings.

See also DYSARTHRIAS: MANAGEMENT.

—Joseph R. Duffy

References


Further Readings


**Dysarthrias: Management**

*Dysarthria* is a collective term for a group of neurologically speech disorders caused by damage to mechanisms of motor control in the central or peripheral nervous system. The dysarthrias vary in nature, depending on the particular neuromotor systems involved. Consequently, a number of issues are considered when devising a management approach for a particular patient. These issues include the type of dysarthria (reflecting the underlying neuromuscular status), the physiological processes involved, severity, and the expected course.

Management of the dysarthrias is generally focused on improving the intelligibility and naturalness of speech, or on helping the speaker convey more communicative intent using speech plus the environment, context, and augmentative aids. *Intelligibility* refers to the degree to which the listener is able to understand the acoustic signal (Kent et al., 1989). *Comprehensibility* refers to the dynamic process by which individuals convey communicative intent, using the acoustic signal plus all information available from the environment (Yorkston, Strand, and Kennedy, 1996). In conversational interaction, listeners take advantage of environmental cues such as facial expression, gestures, the situation, the topic, and so on. As the acoustic speech signal becomes more degraded, contextual information becomes more critical for maintaining comprehensibility.

Decisions regarding whether to focus treatment on intelligibility or on comprehensibility depend largely on the severity of the dysarthria. Management for mildly dysarthric individuals focuses on improving intelligibility and naturalness. Individuals with moderate levels of severity benefit from both intelligibility and comprehensibility approaches. Finally, management of very severe dysarthria often focuses on augmentative communication.

Management focus also depends on whether the dysarthria is associated with a condition in which physiological recovery is likely to occur (e.g., cerebrovascular accident) versus one in which the dysarthria is likely to get progressively worse (e.g., amyotrophic lateral sclerosis [ALS]). For patients with degenerative diseases such as ALS, early treatment may focus on maintaining intelligibility. Later in the disease progression, the focus of treatment is less on the acoustic signal and more on communicative interaction, maximizing listener support and environmental cues, allowing the patient to continue to use speech for a much longer period of time before having to use augmentative and alternative communication. Yorkston (1996) provides a comprehensive review of the treatment efficacy literature for the dysarthrias associated with a number of different neurological disorders.

**Intelligibility**

Deficits in intelligibility vary according to the type of dysarthria as well as the relative contribution of the basic physiological mechanisms involved in speech: respiration, phonation, resonance, and articulation. Medical (e.g., surgical, pharmacological), prosthetic (e.g., palatal lift), and behavioral interventions are used to improve the function of those physiological systems.

Behavioral intervention for respiratory support focuses on achieving and maintaining a consistent subglottal air pressure level, allowing adequate loudness and length of breath groups (Yorkston et al., 1999). Methods to improve respiratory support (Netsell and Daniel, 1979; Hixon, Hawley, and Wilson, 1982) often involve having the speaker blow and maintain target levels of water pressure (i.e., 5 cm H2O) for 5 seconds. Sustained phonation tasks are also used, giving the speaker feedback on maintained loudness. Finally, individuals are encouraged to produce sentences with appropriate
phrase lengths, focusing on maintaining adequate respiratory pressure. In each case, the clinician works to focus the speaker’s attention and effort toward taking in more air and using more force with exhaled air. Occasionally speakers may release too much airflow during speech. Netsell (1995) has suggested the use of inspiratory checking, in which patients are taught to use the inspiratory muscles to counter the elastic recoil forces of the respiratory system. For individuals who exhibit discoordination (e.g., ataxic dysarthria), respiratory treatment is focused on helping the speaker consistently initiate phonation at appropriate inspiratory lung volume levels, taking the next breath at the appropriate phrase boundary, given the expired lung volume level.

Laryngeal system impairment frequently results in either hypophonia, in which the vocal folds do not achieve adequate closure for phonation (as in flaccid dysarthria, or the hypokinetic dysarthria that accompanies Parkinson’s disease), or hyperphonia, in which the vocal folds exhibit too much closure (as in spastic dysarthria). Individuals with lower motor neuron deficits involving the laryngeal muscles may benefit from surgical intervention either to medialize the vocal fold or to augment the bulk of the fold. The most common procedure for medialization is a type I thyroplasty, often with aretynoid abduction (Ishiki, Okamura, and Ishikawa, 1975; Nasseri and Maragos, 2000). Teflon and autogenuous fat are also used to increase the bulk of a paralyzed or atrophied fold (Heikki, 1998). Patients with myasthenia gravis are typically successfully treated with anticholinesterase drugs or with a thyromectomy. This medical management usually results in improvement in their voice and vocal fatigue.

Behavioral treatment for hypophonia focuses on increasing glottal closure, but it also requires that the patient maximize respiratory pressures. For mild weakness, exercises to increase the patient’s awareness of efficient glottal adduction, without extraneous supraglottic tension, are helpful. Effort closure techniques such as pushing and grunting may maximize vocal fold adduction (Rosenbek and LaPoint, 1985). The Lee Silverman Voice Therapy Program (Ramig et al., 1995, 1996) has been shown to be efficacious for individuals with Parkinson’s disease and is a commonly used therapy technique to reduce the hypophonic aspects of their dysarthria.

Treatment of phonation due to laryngeal spasticity is difficult, and behavioral intervention typically is not successful for this group of patients. Techniques to facilitate head and neck relaxation as well as laryngeal relaxation, strategies to maximize efficiency of the respiratory system, and the use of postural control may be helpful. Patients with phonatory deficits due to laryngeal dystonia pose similar problems. Medical management, such as botulinum toxin injection, is frequently used to improve the vocal quality of individuals with spasmodic dysphonia and laryngeal dystonias.

Behavioral approaches to the treatment of resonance problems focus on increasing the strength and function of the soft palate, but researchers and clinicians disagree as to their effectiveness. Kuehn and Wachtel (1994) suggest the use of continuous positive airway pressure in a resistance exercise program to strengthen the velopharyngeal muscles. A common prosthetic approach is to use a palatal lift, which is a rigid appliance that covers the hard palate and extends along the surface of the soft palate, raising it to the pharyngeal wall. Palatal lifts should be considered for patients who are consistently unable to achieve velopharyngeal closure and who have relatively isolated velopharyngeal impairment.

Treatment focused on improving articulation often uses the hierarchical practice of selected syllable, words and phrases (Robertson, 2001). However, because articulatory imprecision may be due to reduced respiratory support, velopharyngeal insufficiency, or rate control, the treatment of articulation is not always focused on improving the place and manner of articulatory contacts. When specific work on improving articulatory function is warranted, behavioral approaches involve focusing the speaker’s attention on increased effort for bigger and stronger movements. Compensatory strategies such as using a different place of articulation or exaggerating selected articulatory movements may be used (DeFao and Schaefer, 1983). The use of minimal contrasts (tie/sigh) or intelligibility drills (having the speaker produce a carefully selected set of stimulus words) focus the speaker’s attention on making specific sound contrasts salient and clear. Strength training is sometimes advocated, but only when the speaker is habitually generating less force than is necessary for speech and has the capacity to increase strength with effort. Strengthening is most appropriate for speakers with flaccid dysarthria; it is contraindicated for patients with disorders such as myasthenia gravis, in which muscular activity causes increasing weakness, and for patients with degenerative disorders such as ALS.

Surgical and medical management may also improve articulation. Neural anastomosis is sometimes used to improve function to a damaged nerve, usually the seventh cranial nerve (Daniel and Guitar, 1978). Botulinum toxin has been used to improve speech in speakers with orofacial and mandibular dystonias (Schulz and Ludlow, 1991). Although pharmacological treatment is frequently used to decrease limb spasticity, its effects on articulation are less clear (Duffy, 1995). Medications to decrease tremor or chorea sometimes help improve speech by reducing the extraneous movement.

Rate control is frequently the focus of treatment for dysarthric individuals. Yorkston et al. (1999) point out that this variable alone may result in the most dramatic changes in speech intelligibility for some individuals. Rate control is most effective for individuals with hypokinetic or ataxic dysarthria, but it may be appropriate for individuals with other types of dysarthria as well. Rate reduction improves intelligibility by facilitating increased precision of movement through the full range, by facilitating more appropriate breath group units, and by allowing listeners more time to process the degraded acoustic signal.
Behavioral approaches are geared toward slowing the rate by increasing consonant and vowel duration, increasing interword interval durations, and increasing pause time at phrasal boundaries, while working to avoid any diminution in speech naturalness. Instrumentation can also be helpful in rate control. In delayed auditory feedback, the speaker’s own voice is fed back to the speaker through earphones after an interval delay. This technique typically slows the rate of speech and improves intelligibility. Visual biofeedback is used to facilitate the speaker’s use of pause time and to slow the rate. Oscilloscopes can provide real-time feedback regarding rate over time. Computer screens can be used that cue the speaker to a target rate and mark the location of pauses (Beukelman, Yorkston, and Tice, 1997).

Comprehensibility

When there is evidence that the individual is able to improve respiratory, phonatory, articulatory, or resonating aspects of speech through behavioral, prosthetic, or medical management, intelligibility is the primary focus of management. However, as the severity of dysarthria increases, management focuses more on the communication interaction between the dysarthric speaker and his or her communicative partners. Management strategies are designed to help the listener maximize the use of context to improve ability to understand even a very degraded acoustic signal. Such strategies include being sure to have the listener’s attention, making eye contact, providing (or asking for) the topic, signaling topic changes, reducing environmental noise, using simple but complete grammatical constructs, using predictable wording, adding gestures if possible, and using alphabet board supplementation. Also important is to adopt a consistent strategy for communication repair that is agreed upon by both speaker and listener. By working on communication interaction between speaker and listener, the dysarthric individual is often able to continue to use speech as a primary mode of communication.

See also DYSARTHRIAS: CHARACTERISTICS AND CLASSIFICATION.

—Edythe A. Strand

References


Further Readings


Dysphagia, Oral and Pharyngeal

"Dysphagia" is an impaired ability to swallow. Dysphagia can result from anatomic variation or neuromuscular impairment anywhere from the lips to the stomach. Although some investigators choose to consider the voluntary oral preparatory stage of deglutition as a separate stage, swallowing is traditionally described as a three-stage event (oral, pharyngeal, and esophageal). Historically, research as well as evaluation and treatment of dysphagia were directed primarily toward the esophageal stage, which is generally treated by a gastroenterologist. However, over the past few decades, speech-language pathologists have become increasingly responsible for the research in, as well as the diagnosis and treatment of, the oral and pharyngeal aspects of deglutition.

The neuroanatomical substrate of dysphagia reflects lower motor neuron innervation by cranial nerves V, VII, IX, X, and XII. Dysphagia can result from unilateral or bilateral cortical insult. Within the cortex, primary sites that contribute to deglutition include the premotor cortex, primary motor cortex, primary somatosensory cortex, insula, and the ventroposterior medial nucleus of the thalamus (Alberts et al., 1992; Daniels, Fountas, Iglesia, et al., 1996; Daniels and Fountas, 1997). Other portions of the cortical system have also been found to be active during swallowing (Hamdy et al., 1999, 2001; Martin et al., 2001).

Dysphagia is associated with an increased risk of developing malnutrition and respiratory complications such as aspiration pneumonia. In a study by Schmidt et al. (1994), the odds ratio that pneumonia would develop was 7.6 times greater for stroke patients who were identified as aspirators than for stroke patients who did not aspirate. Furthermore, the odds ratio of dying was 9.2 times greater for patients who aspirated thickened viscosities than for those who did not aspirate or who aspirated only thin fluids. Davalos et al. (1996) studied the effects of dysphagia on nutritional status in stroke patients who had similar nutritional status at the time of hospital admission. One week after the stroke, 48.3% of patients who developed dysphagia while in the hospi-
patients with dysphagia, 276 (59%) of the 469 patients who aspirated were found to have silent aspiration (Smith et al., 1999). (3) When 47 stroke patients with mixed sites of lesions were examined, 24 (51%) of the patients were found to aspirate; of those 24 patients, 11 (46%) were silent aspirators (Horner and Massey, 1988). (4) In a study of 107 patients in a rehabilitation facility, 43 (40%) were found to aspirate on videofluoroscopic examination; however, clinical evaluation identified only 18 (42%) of the aspirators (Splaingard et al., 1988). Because of the additional expense encountered in caring for patients with respiratory or nutritional complications, studies such as these support the argument that money, as well as life, can be saved when patients are properly evaluated.

Dysphagia can occur at any age across the life span. Among young adults, traumatic brain injury is a not uncommon cause of acquired dysphagia, whereas elderly individuals are more likely to acquire dysphagia as a result of illness. However, young adults are also susceptible to the same causes of dysphagia as the elderly. Neurological disorders take a particular toll: it has been estimated that 300,000–600,000 persons per year experience dysphagia secondary to neurological disorders, and the greatest percentage of these experience dysphagia secondary to stroke (Doggett et al., 2001). After stroke and neurological disease, the most frequent causes of dysphagia in adults include muscle disease, head and neck surgery, radiation to the head and neck, dementia, end-plate disease, traumatic brain injury, systemic disease, cervical spine disease, medication effects, and senescent changes in the sensorimotor system.

Evaluation and Treatment

There are various methods for studying the swallow. The choice of method for a particular patient depends on the information that is sought. When a patient is first seen, the assessment begins with a clinical examination (Perlman et al., 1991). The clinical examination provides important information that assists in the decision-making process, but it is not intended to identify the underlying variables that result in difficulty with oral intake. Furthermore, this examination provides no information relative to the pharyngeal stage of the swallow and does not elicit adequate information to determine proper therapy. Therefore, the clinician will turn to one or more imaging modalities or other specific techniques.

Videofluoroscopy is the most frequently used assessment technique because it provides the most complete body of information. Interpretation of this examination is performed after observation of no less than two dozen events within the oral cavity, pharynx, and larynx. For most patients, this is the only instrumental procedure that will be performed.

Endoscopy permits the examiner to evaluate the status of vocal fold function, the extent of vallecular or pyriform sinus stasis, and the presence of spillover or of delayed initiation of the pharyngeal stage of the swallow (Langmore, Schatz, and Olsen, 1988; Aviv et al., 2000). Additionally, the view of velar function is superior to

Table 1. Clinical Signs Suggestive of Dysphagia in Adults

<table>
<thead>
<tr>
<th>Clinical Signs Suggestive of Dysphagia in Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty triggering the swallow</td>
</tr>
<tr>
<td>Difficulty managing oral secretions, with or without drooling</td>
</tr>
<tr>
<td>Abnormal or absent laryngeal elevation during swallow attempts</td>
</tr>
<tr>
<td>Choking or coughing during or after intake of food or liquid</td>
</tr>
<tr>
<td>Wet-sounding cough</td>
</tr>
<tr>
<td>Wet, gurgly voice quality</td>
</tr>
<tr>
<td>Decreased oral sensation</td>
</tr>
<tr>
<td>Weak sign of the cough</td>
</tr>
<tr>
<td>Prolonged oral preparation with food</td>
</tr>
<tr>
<td>Inability to clear the mouth of food after intake</td>
</tr>
<tr>
<td>Absent gag reflex</td>
</tr>
<tr>
<td>Food or liquid leaking from a tracheostomy site</td>
</tr>
<tr>
<td>Fullness or tightness in the throat (globus sensation)</td>
</tr>
<tr>
<td>Food or liquid leaking from the nose</td>
</tr>
<tr>
<td>Regurgitation of food</td>
</tr>
<tr>
<td>Sensation of food sticking in the throat or sternal region</td>
</tr>
<tr>
<td>Xerostomia (dry mouth)</td>
</tr>
<tr>
<td>Odynophagia (pain on swallowing)</td>
</tr>
<tr>
<td>Repeated incidents of upper respiratory infections with or without a diagnosis of aspiration pneumonia</td>
</tr>
<tr>
<td>Tightness or pain in the chest, particularly after eating or when lying down</td>
</tr>
<tr>
<td>Heartburn or indigestion</td>
</tr>
<tr>
<td>Unintended weight loss not related to disease</td>
</tr>
</tbody>
</table>

Inadequate nutrition negatively affects the ability of the immune system to fight disease and contributes to the development of respiratory and cardiac insufficiency, the formation of decubitus ulcers, and impaired gastrointestinal function. The already comprised patient can become increasingly comprised, which prolongs the hospital length of stay and increases medical costs.

Certain clinical signs help to alert health care providers to the likely presence of dysphagia. Table 1 lists commonly observed clinical signs that are suggestive of dysphagia in the adult population. The absence of any or all of these signs does not indicate that a patient has a safe swallow or that the patient is able to ingest an adequate number of calories by mouth to remain properly nourished. For example, a diminished or absent gag has not been found to distinguish aspirators from non-aspirators (Horner and Massey, 1988). Many of these signs can be indicative of a serious medical illness. Therefore, patients who exhibit these signs and who have not been seen by a physician should be referred for medical examination.

Although clinical indicators have been found to have a relationship to laryngeal penetration, a significant number of patients who aspirate do so with no clinical indication. The incidence of silent aspiration is very high, and the difficulty of detecting it is suggested by the following: (1) Discriminant analysis of 11 clinical indicators resulted in identification of the presence of aspiration in only 66% of patients (Linden, Kuhlemeier, and Patterson, 1993). (2) In a heterogeneous group of 1101
that obtained with videofluoroscopy. Information relating to the oral stage of deglutition, the extent of elevation of the hyoid bone, or information on the larynx or pharynx during the moment of swallow is not observed with endoscopy.

Ultrasound allows for observation of the motion of the tongue (Sonies, 1991). Additionally, the shadow reflected from the hyoid bone permits the examiner to observe and to measure the displacement of the hyoid. The advantages to using ultrasound for assessing the oral stage of the swallow are the absence of exposure to ionizing radiation and the fact that the parent can hold an infant or small child and feed the child a familiar food while the examination is being performed. The information obtainable with ultrasound is restricted to the oral stage of deglutition. When a small child has a tracheotomy tube, it is often extremely difficult to obtain a good ultrasound image, because the tracheostomy tube prohibits good transducer placement.

Muscle paralysis is best determined with intramuscular electromyography (Cooper and Perlman, 1997; Perlman et al., 1999). In the examination of swallowing, it is advisable to use bipolar hooked wire electrodes, because needle electrodes can cause discomfort and the subject may alter the swallowing pattern.

Respirodeglutometry (RDG) is a method for assessing the coordination of respiration and deglutition (Perlman, Ettema, and Barkmeier, 2000). This technique is presently being investigated to determine the physiological correlates of RDG output and to determine changes in the respiratory-swallowing pattern as a function of age and various medical diagnoses.

Decisions regarding behavioral, medical, or surgical intervention are made once the evaluation has been completed. Therapeutic intervention is determined as a function of the anatomical and physiological observations that were made during the evaluation process. Specific treatments are beyond the scope of this discussion but can be found in textbooks listed in Further Readings.

—Adrienne L. Perlman

References


Further Readings


Early Recurrent Otitis Media and Speech Development

Otitis media can be defined as inflammation of the middle ear mucosa, resulting from an infectious process (Scheidt and Kavanagh, 1986). When the inflammation results in the secretion of effusion, or liquid, into the middle ear cavity, the terms otitis media with effusion (OME) and middle ear effusion (MEE) are often used. Middle ear effusion may be present during the period of acute inflammation (when it is known as acute otitis media), and it may persist for some time after the acute inflammation has subsided (Bluestone and Klein, 1996a).

The prevalence of OME in young children is remarkably high. OME has been described as one of the most common infectious diseases of childhood (Bluestone and Klein, 1996a) and evidence from several large studies supports this conclusion. For example, in a prospective epidemiologic study of 2253 children enrolled by age 2 months, Paradise et al. (1997) reported that nearly 80% of children had at least one episode of OME by age 12 months of age; more than 90% had an episode by age 24 months. Low socioeconomic status, male sex, and amount of exposure to other children were associated with an increased prevalence of OME during the first 2 years of life (Paradise et al., 1997).

The literature addressing the hypothesis that early recurrent OME poses a threat to children’s speech and language development is large and contentious (for reviews, see Stool et al., 1994; Shriberg, Flipsen, et al., 2000). OME has been reported to result in adverse effects, no effects, and small beneficial effects (e.g., Shriberg, Friel-Patti, et al., 2000), sometimes within the same study. Substantive methodological differences may account for much of the disparity in findings, with the method by which OME is diagnosed in different studies being critically important. The gold standard for diagnosing OME is an examination of the tympanic membrane via pneumatic otoscopy, after any necessary removal of cerumen, to determine whether indicators of effusion such as bulging, retraction, bubbling, or abnormal mobility are present (Stool et al., 1994; Bluestone and Klein, 1996b). Behavioral symptoms such as irritability are neither sensitive nor specific to the condition, and the validity of parental judgments concerning the frequency or duration of episodes of OME is poor even when repeated feedback is provided (Anteunis et al., 1999). In addition, it is important that OME be documented prospectively rather than via retrospective reports or chart reviews, because a substantial percentage of apparently healthy and symptom-free children are found to have OME on otoscopic assessment (Bluestone and Klein, 1996a).

A second important difference among studies of OME and speech development is the extent to which hearing levels are documented. The hypothesis that OME poses a threat to speech or language development is typically linked to the assumption that effusion causes conductive hearing loss, which prevents children from perceiving and processing speech input in the usual fashion (e.g., K. Roberts, 1997). However, the presence of effusion is a poor predictor of hearing loss. Although hearing thresholds for the majority of children with MEE fall between 21 and 30 dB (mild to moderate degrees of impairment), thresholds from 0 to 50 dB are not uncommon (Bess, 1986). Hearing thresholds must be measured directly to determine whether OME has effects on development independent of its variable effects on hearing (e.g., Shriberg, Friel-Patti, et al., 2000).

Studies also vary with respect to their ability to separate the contribution of OME to poor developmental outcome from the effects of other variables with which OME is known to be associated, such as sex and socioeconomic status. As noted earlier, OME is significantly more prevalent in males than in females, and in children from less privileged backgrounds than in their more privileged counterparts (Paradise et al., 1997; Peters et al., 1997). Statistical procedures are necessary to control for such confounding in order to distinguish the effects of OME from those of other variables. Several recent studies have shown that after controlling for socioeconomic confounds, OME accounts for little if any of the variance in developmental outcome measures (e.g., J. E. Roberts et al., 1998; Paradise et al., 2000).

Finally, studies have also differed substantially in the measures used to document the outcome variable of speech development and in the extent to which effect sizes for significant differences on outcome measures are reported (cf. Casby, 2001). No accepted standard metric for speech delay or disorder currently exists, although the Speech Disorders Classification System developed by...
Shriberg et al. (1997) represents an important advance toward meeting this need. Instead, the effects of OME on speech have been sought on a wide range of articulatory and phonological measures, not all of which are known to be predictive of eventual speech outcome (e.g., Rvachew et al., 1999).

When these cautions are borne in mind, the literature on early recurrent otitis media and speech development suggests converging evidence that OME in and of itself has a negligible relationship to early speech development. Several prospective investigations have shown little or no relationship between cumulative duration of otitis media (documented otoscopically) and measures of speech production in otherwise healthy children. In a longitudinal study of 55 low-SES children, J. E. Roberts et al. (1988) found no correlation between OME and number of consonant errors or phonological processes on a single-word test of articulation at ages 3, 4, 5, 6, 7, or 8 years. Shriberg, Friel-Patti, et al. (2000) examined ten speech measures derived from spontaneous speech samples obtained from 70 otherwise healthy, middle to upper middle class 3-year-olds who were classified according to the number of episodes of OME from 6 to 18 months of age; only one significant speech difference was found, in which the group with more OME paradoxically obtained higher intelligibility scores than the group with fewer bouts of OME. Paradise et al. (2000) likewise found no relationship between cumulative duration of MEE and scores on the Percentage of Consonants Correct–Revised measure (PCC-R; Shriberg, 1993) in 241 sociodemographically diverse children at age 3 years. Paradise et al. (2001) reported that PCC-R scores from children with even more persistent MEE from 2 to 36 months of age did not differ significantly from those of children with the less persistent levels of effusion reported by Paradise et al. (2000). Further, children with persistent and substantial MEE who were randomly assigned to undergo prompt tympanostomy tube placement had no better PCC-R scores at age 3 than children who underwent tube placement after a delay of 6–9 months, during which their MEE persisted (Paradise et al., 2001). These findings of little or no relationship between OME and speech development mirror those of several recent reports showing negligible associations between early OME and later oral and written language performance (Peters et al., 1997; Casby, 2001).

By contrast with these negative findings concerning the impact of OME, several studies in which hearing was documented showed poorer speech outcomes for children with elevated hearing thresholds. In a sample of 70 middle to upper middle class 3-year-olds who received otoscopic evaluations every 6 weeks and hearing evaluations every 6 months between 6 and 18 months of age, Shriberg, Friel-Patti, et al. (2000) reported that children with hearing loss, defined as average thresholds ≥20 dB (HL) during one evaluation between 6 and 18 months of age, had a significantly increased risk of scoring more than 1.3 standard deviations below the sample mean on several percentage-consonants-correct metrics. Shriberg et al. note the need for some caution in interpreting these findings, given that increased risk was not found across all speech metrics and that confidence intervals for risk estimates were wide. In addition, the results of structural equation modeling suggested that hearing loss did not operate directly to lower speech performance, but rather was mediated significantly by language performance, providing another indication of the need for multifactorial approaches to identifying the factors and pathways involved in normal and abnormal speech development.

Although the best available current evidence suggests that OME itself does not represent a significant risk to speech development in otherwise healthy children, the question of whether OME may contribute independently to outcome when it occurs in conjunction with other risk factors or health conditions (e.g., Wallace et al., 1996; J. E. Roberts et al., 1998; Shriberg, Flipsen, et al., 2000) remains open. Additional investigations that include prospective otoscopic diagnosis of OME; frequent and independent assessments of hearing; valid, reliable assessments of both medical and sociodemographic risk factors; and a multifactorial analytic strategy will be needed to answer this question.

See also OTITIS MEDIA: EFFECTS ON CHILDREN’S LANGUAGE.

—Christine Dollaghan and Thomas Campbell

References


risk factors during the first two years of life. Pediatrics, 9, 318–333.


Further Readings


Laryngectomy

Total laryngectomy is a surgical procedure to remove the larynx. Located in the neck, where it is commonly referred to as the Adam’s apple, the larynx contains the vocal folds for production of voice for speech. Additionally, the larynx serves as a valve during swallowing to prevent food and liquids from entering the airway and lungs. When a total laryngectomy is performed, the patient loses his or her voice and must breathe through an opening created in the neck called a tracheostoma.

Total laryngectomy is usually performed to remove advanced cancers of the larynx, most of which arise from prolonged smoking or a combination of tobacco use and alcohol consumption. Laryngeal cancers account for less than 1% of all cancers. About 10,000 new cases of laryngeal cancer are diagnosed each year in the United States, with a male-female ratio approximately 4 to 1 (American Cancer Society, 2000). The Surveillance, Epidemiology and End Results (SEER) program of the National Cancer Institute (Ries et al., 2000) reports that laryngeal cancer rates rise sharply in the fifth, sixth, and first half of the seventh decades of life (Casper and Colton, 1998). The typical person diagnosed with cancer of the larynx is a 60-year-old man who is a heavy smoker with moderate to heavy alcohol intake (Casper and Colton, 1998). Symptoms of laryngeal cancer vary, depending on the exact site of the disease, but persistent hoarseness is common. Other signs include lowered pitch, sore throat, a lump in the throat, a lump in the neck, earache, difficulty swallowing, coughing, difficulty breathing, and audible breathing (National Cancer Institute, 1995). It is estimated that there are 50,000 laryngectomees (laryngectomized people) living in the United States today.

As a treatment of laryngeal cancer, total laryngectomy is a proven technique to control disease. The primary disadvantages of total laryngectomy are the loss of the vocal folds that produce voice for speech and the need for a permanent tracheostomy for breathing. Before the introduction of the extended partial laryngectomy, patients with cancer of the larynx were treated primarily with total laryngectomy (Weber, 1998). Today, early and intermediate laryngeal cancers can be cured with conservation operations that preserve voice, swallowing, and nasal breathing, and total laryngectomy is performed only in cases of very advanced cancers that are bilateral, extensive, and deeply invasive (Pearson, 1998). Radiation therapy is often administered before or
after total laryngectomy. In addition, radiation therapy alone and sometimes in combination with chemotherapy has proved to be curative treatment for laryngeal cancer, depending on the site and stage of the disease (Chyle, 1998). Controversy and research continue over nonsurgical versus surgical intervention or a combination of these for advanced laryngeal cancer, weighing the issues of survival, preservation of function, and quality of life (Weber, 1998).

A person with laryngeal cancer and the family members have many questions about survival, treatment options, and the long-term consequences and outcomes of various treatments. An otolaryngologist is the physician who usually diagnoses cancer of the larynx and provides information about possible surgical interventions. A radiation oncologist is a physician consulted for opinions about radiation and chemotherapy approaches to management. If the patient decides to have a total laryngectomy, a speech pathologist meets with the patient and family before the operation to provide information on basic anatomy and physiology of normal breathing, swallowing, and speaking, and how these will change after removal of the larynx (Keith, 1995). Also, the patient is informed that, after a period of recovery and rehabilitation, and with a few modifications, most laryngectomized people return to the same vocational, home, and recreational activities they participated in prior to the laryngectomy.

Besides voicelessness, laryngectomized persons experience other changes. Since the nasal and oral tracts filter the air as well as provide moisture and warmth in normal breathing, laryngectomized people often require an environment with increased humidity, and they may wear heat- and moisture-exchanging filters over their tracheostoma to replicate the functions of nasal and oral breathing (Grolman and Schouwenberg, 1998). There is no concern for aspiration of food and liquids into the lungs after total laryngectomy, because the respiratory and digestive tracts are completely separated and no longer share the pharynx as a common tract. Unless the tongue is surgically altered or extensive pharyngeal or esophageal reconstruction beyond total laryngectomy is performed, most laryngectomized people return to a normal diet and have few complaints about swallowing other than that it may require additional effort (Logemann, Paukloski, and Rademaker, 1997).

There are nonspeech methods of communicating that can be used immediately after total laryngectomy. These include writing on paper or on a slate, pointing to letters or words or pictures on a speech or communication board, gesturing with pantomimes that are universally recognizable, using e-mail, typing on portable keyboards or speech-generating devices, and using life-line emergency telephone monitoring systems. None of these methods of communicating is as efficient or as personal as one’s own speech.

A common fear of the laryngectomee is that he or she will never be able to speak without vocal folds. There are several methods of alaryngeal (without a larynx) speech. Immediately or soon after surgery, a laryngectomized person can make speech movements with the tongue and lips as before the surgery, but without voice. This silent speech is commonly referred to as “mouthing words” or “whispering”; however, unlike a normal whisper, air from the lungs does not move through the mouth after a laryngectomy. The effectiveness of the technique is variable and depends largely on the laryngectomee’s ability to precisely articulate speech movements and the ability of others to recognize or “read” them.

Artificial larynges have been used since the first recorded laryngectomy in 1873 (Billroth and Gussenbauer, 1874). Speech with an artificial larynx, also known as an electrolarynx, can be an effective method of communicating after laryngectomy, and many people can use one of these instruments as early as a day or two after surgery. Most modern instruments are battery powered and produce a mechanical tone. Usually the device is pressed against the neck or under the chin at a location where it produces the best sound, and the person articulates this “voice” into speech. If the neck is too swollen after surgery or the skin is hard as a result of radiation therapy, the tone of the artificial larynx may not be conducted into the throat sufficiently for production of speech. In this circumstance it may be possible to use an oral artificial larynx with a plastic tube to place the tone directly into the mouth, where it is articulated into speech. Speech with an artificial larynx has a sound quality that is mechanical, yet a person who uses an artificial larynx well can produce intelligible speech in practically all communication situations, including over the telephone. Most laryngectomized people require training by a speech pathologist to use an artificial larynx optimally.

A laryngectomized person may be able to learn to use esophageal voice, also known as esophageal speech. For this method, commonly known as “burp speech,” the person learns to use the esophagus (food tube) to produce voice. First the laryngectomee pumps or sucks air into the esophagus. Sound or “voice” is generated as the air trapped in the upper esophagus moves back up through the narrow junction of the pharynx and esophagus known as the PE segment. Then the voice is articulated into speech by the tongue and lips.

A key to producing successful esophageal speech is getting air into the esophagus consistently and efficiently, followed by immediate sound production for speech. Esophageal speech has distinct advantages over other alaryngeal speech techniques. The esophageal speaker requires no special equipment or devices, and the speaker’s hands are not monopolized during conversation. A significant disadvantage of esophageal speech is that it takes a relatively long time to learn to produce voice that is adequate for everyday speech purposes. Additionally, insufficient loudness and a speaking rate that is usually slower than before laryngectomy are common concerns of esophageal speakers. Although some become excellent esophageal speakers, many do not attain a level of fluent speech sufficient for all communicative situations.
Tracheoesophageal puncture with a voice prosthesis is another method of alaryngeal voice production (Singer and Blom, 1980; Blom, 1998). During the surgery to remove the larynx or at a later time, the surgeon makes an opening (puncture) just inside and inferior to the superior edge of the tracheostoma. The opening is a tract through the posterior wall of the trachea and the anterior wall of the esophagus. Usually a catheter is placed in the opening and a prosthesis is placed a few days later. A speech pathologist specially trained in tracheoesophageal voice restoration measures the length of the tract between the trachea and the esophagus, and a silicone tube with a one-way valve—a voice prosthesis—is placed in the puncture site. The prosthesis is nonpermanent and must be replaced periodically. It does not generate voice itself. When the person exhales and the tracheostoma is covered with a thumb, finger, or special valve, air from the lungs moves up the trachea, through the prosthesis, into the upper esophagus, and through the PE segment to produce voice. Because lung air is used to produce voice with a tracheoesophageal puncture, the speech characteristics of pitch and loudness, rate, phrasing, and timing more closely resemble the laryngectomee’s presurgical speech qualities than can be achieved with other forms of alaryngeal speech. For many laryngectomees, fluent speech can be achieved soon after placement of a voice prosthesis.

There are disadvantages associated with tracheoesophageal puncture. The laryngectomee may dislike using a thumb or finger to cover the tracheostoma when speaking, and use of a tracheostoma valve for hands-free speaking may not be possible. Expenses associated with tracheoesophageal puncture include those for initial training in the use and maintenance of the prosthesis with a speech pathologist and subsequent clinical visits for modification or replacement of the voice prosthesis, and ongoing costs of prosthesis-related supplies. If the PE segment is hypertonic and the tracheoesophageal voice is not satisfactorily fluent for conversation, or if it requires considerable effort to produce, injection of botulinum neurotoxin, commonly known as Botox, may be required (Hoffman and McCulloch, 1998; Lewin et al., 2001), or myotomy of the pharyngeal constrictor muscles may be considered (Hamaker and Chessman, 1998).

Historically, laryngectomees and speech pathologists have felt strongly about one form of alaryngeal speech being superior to others. In the 1960s, newly laryngectomized persons were discouraged from using artificial larynges, which were thought to delay or interfere with the learning of esophageal speech (Lauder, 1968). Today some think tracheoesophageal speech is superior because many laryngectomees are able to speak fluently and fairly naturally with this method only a few weeks after surgery. Others maintain that esophageal speech, with no reliance on a prosthesis or other devices, is the gold standard against which all other methods should be compared (Stone, 1998). Most believe any form of speech after laryngectomy is acceptable and should be encouraged, since speaking is a fundamental and essential part of being human.

People who undergo total laryngectomy experience the same emotions of shock, fear, stress, loss, depression, and grief as others with life-threatening illnesses. Along with regular medical follow-up to monitor for possible recurrence of cancer and to review all the body systems, laryngectomized persons may benefit from referral to other professionals and resources for psychological, marital, nutritional, rehabilitation, and financial concerns.

The International Association of Laryngectomees and the American Cancer Society provide services to laryngectomized persons. They sponsor peer support groups, provide speech therapy, and distribute educational materials on topics of interest, such as cardiopulmonary resuscitation for neck breathers, smoking cessation, and specialized products and equipment for laryngectomized persons.

See also ALARYNGEAL VOICE AND SPEECH REHABILITATION.

—Jack E. Thomas

References
Mental Retardation and Speech in Children

Mental retardation is defined by the American Association on Mental Retardation as significantly subaverage intellectual functions with related limitations in social and behavioral skills. According to the most recent estimates (Larson et al., 2001), the prevalence of mental retardation in the noninstitutionalized population of the United States is 7.8 people per thousand; if institutionalized individuals are included in the prevalence rates, the number increases to 8.73 per thousand. Mental retardation is associated with limitations in learning and in the ability to communicate, and has a profound effect on a child’s ability to learn to talk. At one time it was believed that the language acquisition of all persons with mental retardation represented a slow-motion version of normal language development. This hypothesis has two major flaws: first, patterns of language development vary across types of mental retardation, and second, within a single type of mental retardation, there is considerable heterogeneity.

The majority of research on children with mental retardation has involved children with Down syndrome (or trisomy 21). This syndrome is the most common genetic cause of mental retardation, occurring in approximately one out of every 800 births. Because Down syndrome is identifiable at birth, researchers have been able to trace developmental patterns from the first months of life. The development of speech and language is severely affected in children with Down syndrome, with levels lower than would be expected, given mental age (Miller, 1988). Speech intelligibility is compromised throughout the life span because of problems with articulation, prosody, and voice.

Children with Down syndrome differ from the normal population in respect to a variety of anatomical and physiological features that may affect speech production. These features include differences in the vocal cords, the presence of a high palatal vault and a larger than normal tongue in relation to the oral cavity, weak facial muscles, and general hypotonicity. Although the precise effect of these differences is difficult to determine, they undoubtedly influence speech-motor development and thus the articulatory and phonatory abilities of children with Down syndrome. An additional factor affecting the speech of children with Down syndrome is fluctuating hearing loss associated with otitis media and middle ear pathologies.

Fragile X syndrome, the most common known cause of inherited mental retardation (Down syndrome is more common but is not inherited), has an estimated prevalence of approximately one per 1250 in males and one per 2500 in females, with males exhibiting more severe effects. Little research has been done on the speech of young children with fragile X syndrome. Available reports on older children (Abbeduto and Hagerman, 1997) indicate abnormalities in articulatory development, disflucts, and the presence of anypical rate and rhythm. These abnormalities may be attributed, in part, to differences in the structure and function of the oral-motor systems of boys with fragile X syndrome, including excessive drooling, hypotonia involving the oral-facial muscles, and the presence of a narrow, high-arched palate. Like their peers with Down syndrome, children with fragile X syndrome have a high incidence of otitis media and intermittent hearing loss.

Autism is a developmental disorder with prevalence estimates ranging from two to five per 10,000 (3:1 males). This disorder is characterized by deficits in social
interaction, communication, and play; two out of three children with autism are mentally retarded (Pennington and Bennetto, 1998). Although in phonetic form, the prelinguistic vocalizations are like those of nonretarded infants, social communication skills in the prelinguistic period are atypical. About 50% of autistic children fail to develop spoken language; the other 50% exhibit delays in acquiring language, although not to the same extent as children with Down syndrome do. Speech production is characterized by echolalia and abnormal prosody (see AUTISM).

Williams syndrome, a genetic disorder that includes mental retardation, is relatively rare, occurring in one in 25,000 live births. One of the most striking aspects of Williams syndrome is that, in spite of marked impairments in cognition, linguistic skills appear to be relatively normal (Bellugi, Lai, and Wang, 1997; Mervis and Bertrand, 1997). This dissociation of language and cognition underscores the importance of examining the relationship between mental retardation and speech in a variety of mentally retarded populations.

The foundations for speech development are laid in the first year of life, with the emergence of nonmeaningful vocal types that serve as precursors for the production of words and phrases. Of particular importance is the production of consonant-vowel syllables, such as [baba], which generally appear around age 6–7 months. Phonetically, these “canonical” babbles are similar or even identical to the forms used in first words; thus, the production [mama] may be a nonmeaningful babble at 8 months and a word at 14 months. The difference is recognition of the sound-meaning relationships that are the basis for words. In general, prelinguistic vocal development of infants with mental retardation resembles that of their nonretarded peers in terms of types of vocalizations and schedule for emergence. Infants with retardation begin to produce canonical babble within the normal time frame or with minor delays.

Despite the nearly normal onset of canonical babble, however, the emergence of words is often delayed among infants with mental retardation, particularly those with Down syndrome (Stoel-Gammon, 1997). Research suggests great variability among children in this domain, with a few reports of word use in the second year of life for a few children with Down syndrome but the majority showing first words appearing between 30 and 60 months. The magnitude of the delay cannot be easily predicted from the degree of retardation. Moreover, once words appear, vocabulary growth is relatively slow. Whereas nonretarded children have a vocabulary of 250 words at 24 months, this milestone is not reached until the age of 4–6 years for most children with Down syndrome.

In terms of phonemic development, acquisition patterns for children with mental retardation are similar to those documented for nonretarded children (Rondal and Edwards, 1997). In the early stages, words are “simplified” in terms of their structure: consonant clusters are reduced to single consonants, unstressed syllables are deleted, and consonants at the ends of words may be omitted. Phonemes that are later-acquired in normal populations, primarily fricatives, affricates, and liquids, also pose difficulties for children with mental retardation. Among nonretarded children acquiring English, correct pronunciation of all phonemes is achieved by the age of 8 years. Some reports suggest that the phonologies of children with Williams syndrome may be relatively adultlike by the (chronological) age of 8.

In contrast, individuals with Down syndrome, even when they have a mental age of 8, exhibit many articulation errors. Moreover, comparisons of phonological development in three populations matched for mental age, Down syndrome, non-Down syndrome with mental retardation, and typically developing, revealed a greater number and variety of error types in the children with Down syndrome (Dodd, 1976). A persistent problem in children with Down syndrome, is that their speech is hard to understand (Kumin, 1994). Parents report low levels of intelligibility through adolescence as a result of speech sound errors, rate of speech, disfluencies, abnormal voice quality, and unusual voice quality. There is some indication that children with fragile X syndrome also suffer from low levels of intelligibility (Abbeduto and Hagerman, 1997).

For many children with mental retardation, delays in the acquisition of speech and language may serve as the first indication of a cognitive delay (except for Down syndrome, which is easily diagnosed at birth). Parents may be the first to raise concerns about atypical patterns of development, and pediatricians and social workers should be aware of the link between linguistic and cognitive development. Once mental retardation has been confirmed, assessment typically adheres to traditional practices in speech-language pathology. In the prelinguistic period, which may be quite protracted for some children, assessment is initially based on unstructured observations and parental report. If language is slow to emerge, it is important to assess hearing and oral-motor function. More formal assessments are done in two ways: by means of standardized tests that focus on the individual sounds and structures of a predetermined set of words (i.e., a normed articulation test) and by analyzing samples of conversational speech to determine intelligibility and overall speech characteristics.

Recommendations for the treatment of speech deficits in children with mental retardation range from intervention directed toward underlying causes such as hearing loss and deficits in speech-motor skills (Yalert, 1980) to programs aimed at modifying parent-to-child speech in order to provide optimal input in the face of delayed language acquisition. Most phonological interventions focus on increasing the phonetic repertoire and reducing the number of errors, using therapy techniques similar to those for children with phonological delay or disorder. In some cases, therapy may occur at home, with the parents, as well as in the clinic (Dodd and Leahy, 1989; Cholmain, 1994).

See also COMMUNICATION SKILLS OF PEOPLE WITH DOWN SYNDROME; MENTAL RETARDATION.

—Carol Stoel-Gammon
includes concomitant language deficits. This category is
absence of obvious neuromotor causes and often
children's speech production deficits are predominantly
Motor speech involvement of unknown origin is a rela-
Motor Speech Involvement in Children

Motor speech involvement of unknown origin is a rela-
tively new diagnostic category that is applied when
children's speech production deficits are predominantly
linked to sensorimotor planning, programming, or exec-
ution (Caruso and Strand, 1999). The disorder occurs in
the absence of obvious neuromotor causes and often
includes concomitant language deficits. This category is
broader than and encompasses that of developmental
apraxia of speech (DAS), which refers specifically to
impaired planning, or praxis. Classically, developmental
speech production deficits have been categorized as ei-
ther phonological or DAS. However, recent empirical
evidence suggests that a wider range of children (e.g.,
those with specific language impairment [SLI] or inconsis-
tent speech errors) may exhibit deficits that are influ-
enced by motor variables and, in these cases, may be
classified as motor speech involved.

Although the underlying causes of motor speech
involvement are unclear, there is general evidence that
motor and cognitive deficits often co-occur (Diamond,
2000). Neurophysiological findings support the inter-
action of cognitive and motor development, most nota-
ably in common brain mechanisms in the lateral
perisylvian cortex, the neocerebellum, and the dorso-
lateral prefrontal cortex (Diamond, 2000; Hill, 2001).
Apparently, speech motor and language domains co-
develop and mutually influence one another across
development.

In late infancy, basic movement patterns observed in
babbling are linked to emerging intents and words
(de Boysson-Bardis and Vihman, 1991; Levelt, Roelofs,
and Meyer, 1999). At this level, it is apparent how lan-
guage and motor levels constrain one another. How-
ever, the relations between language and motor levels in
later periods of development have not been specified.
Language models include categories such as concepts,
semantics, syntax, and phonology (Levelt, Roelofs,
and Meyer, 1999). Motor systems are discussed in the very
different terms of cortical inputs to pattern generators in
the brainstem, which in turn provide inputs to motor
neuron pools for the generation of muscle activity. Sen-
sory feedback is also a necessary component of motor
systems (A. Smith, Goffman, and Stark, 1995). Although
it is established that motor and language domains both
show a protracted developmental time course, speech
production models are not explicit about the nature of the
linkages. The general view is that increasingly com-
plex linguistic structures are linked to increasingly com-
plex movements in the course of development. Motor
speech deficits occur when movement variables interfere
with the acquisition of speech and language production.

A large range of speech and language characteristics
have been reported in children diagnosed with motor
speech disorders. In the following summary, emphasis
is placed on those that are at least partially motor in
origin.

Variability. Children with motor speech disorders
have been reported to produce highly variable errors,
even across multiple productions of the same word
(Davis, Jakielski, and Marquardt, 1998). When the defi-
cit involves movement planning, imitation and repeti-
tion may not aid performance (Bradford and Dodd,
1996). Although variability is observed in speech motor
(A. Smith and Goffman, 1998) and phonetic output of
young children who are normally developing, it is ex-
treme and persistent in disordered children. Usually,
variability is discussed as a phonetic error type. How-

References
Abbeduto, L., and Hagerman, R. J. (1997). Language and
communication in fragile X syndrome. Mental Retardation
and Developmental Disabilities Research Reviews, 3, 313–
322.
nication, and neural systems in Williams syndrome.
Mental Retardation and Developmental Disabilities Research
Reviews, 3, 334–342.
Cholmain, C. N. (1994). Working on phonology with young
children with Down syndrome. Journal of Clinical Speech
and Language Studies, 1, 14–35.
Dodd, B. J. (1976). A comparison of the phonological systems
of mental age matched normal, severely abnormal and
Down's syndrome children. British Journal of Disorders of
Communication, 11, 27–42.
Dodd, B. J., and Leahy, P. (1989). Phonological disorders and
mental handicap. In M. Beveridge, G. Conti-Ramsden, and
I. Leudar (Eds.), Language and communication in mentally
handicapped people (pp. 33–56). London: Chapman and
Hall.
Down syndrome in natural settings: Parents' perspective.
Perceptual and Motor Skills, 78, 307–313.
tardation and developmental disabilities: Estimates from
the 1994/1995 National Health Interview Survey Disability
Supplements. American Journal on Mental Retardation
106, 231–252.
Mervis, C. B., and Bertrand, J. (1997). Developmental rela-
tions between cognition and language: Evidence from Wil-
liams syndrome. In L. B. Adamson and M. A. Romski
(Eds.), Communication and language acquisition: Discoveries
from atypical development (pp. 75–106). Baltimore: Paul H.
Brookes.
Miller, J. F. (1998). The developmental asynchrony of lan-
guage development in children with down syndrome. In
L. Nadel (Ed.), The psychobiology of Down syndrome (pp.
Pennington, B. F., and Bennetto, L. (1998). Toward a neuro-
psychology of mental retardation. In J. A. Burack, R. M.
Hodapp, and E. Zigler (Eds.), Handbook of mental retarda-
tion and development (pp. 80–114). Cambridge, UK:
Cambridge University Press.
Rondal, J., and Edwards, S. (1997). Language in mental retar-
syndrome. Mental Retardation and Developmental Dis-
abilities Research Reviews, 3, 300–306.
Yarter, B. H. (1980). Speech and language programs for the
Down's population. Seminars in Speech, Language and
Hearing, 1, 49–61.
ever, kinematic analysis of lip and jaw movement reveals that children with SLI show movement output that is less stable than that of their normally developing peers, even when producing an accurate phonetic segment (Goffman, 1999). Thus, both phonological and motor factors may contribute. Deficits in planning and implementing spatially and temporally organized movements may influence the acquisition of stable phonological units (Hall, Jordan, and Robin, 1993).

**Duration.** Increased movement durations are a hallmark of immature motor systems (B. L. Smith, 1978; Kent and Fornier, 1980). In children with motor speech involvement, the slow implementation of movement may lead to decreased performance on a nonlinguistic diadochokinetic task (Crary, 1993) as well as increased error rates on longer and more complex utterances. An additional error type that may also be related to timing is poor movement coordination across speech subsystems. Such timing deficits in articulatory and laryngeal coordination may lead to voicing and nasality errors. Hence, these errors may have origins in movement planning and implementation. A decreased speech rate provides the child with time to process, plan, and implement movement (Hall, Jordan, and Robin, 1993), but it may also negatively influence speech motor performance.

**Phonetic Movement Organization and Sequencing.** As they develop, children produce increasingly differentiated speech movements, both within and across articulatory, laryngeal, and respiratory subsystems (Gibbon, 1999; Moore, 2001). A lack of differentiated and coordinated movement leads to a collapsing of phonetic distinctions. It follows that segmental and syllabic inventories are reduced for children with motor speech deficits (Davis, Jakielski, and Marquardt, 1998). Vowel and consonant errors may be considered in reference to articulatory complexity. Vowel production requires highly specified movements of the tongue and jaw (Pollock and Hall, 1991). Consonant sounds that are early-developing and that are most frequently seen in the phonetic inventories of children with motor speech deficits make relatively few demands on the motor system (Hall, Jordan, and Robin, 1993). Kent (1992) suggests that early-developing stop consonants such as [b] and [d] are produced with rapid, ballistic movements. Fricatives require fine force control and are acquired later. Liquids, which require highly controlled tongue movements, are learned quite late in the developmental process. Using electropalatography, Gibbon (1999) has provided direct evidence that children with speech deficits contact the entire palate with the tongue, not just the anterior region, in their production of alveolar consonants. Such data indicate that motor control of differentiated tongue movements has not developed in these children. Overall, as proposed by Kent (1992), motor variables account for many aspects of the developmental sequence frequently reported in speech- and language-impaired children.

Syllable shapes may also be influenced by motor factors. The earliest consonant-vowel structure seen in babbling is hypothesized to consist of jaw oscillation without independent control of the lips and tongue (MacNeilage and Davis, 2000). More complex syllable structures probably require increased movement control, such as the homing movement for final consonant production (Kent, 1992).

**Prosodic Movement Organization and Sequencing.** One major aspect of motor development that has been emphasized in motor speech disorders is rhythmicity. Rhythmicity is thought to have origins in prelinguistic babbling (and, perhaps, in early stereotypic movements, such as kicking and banging objects) (e.g., Thelen and Smith, 1994). Rhythmicity underlies the prosodic structure of speech, which is used to convey word and sentence meaning as well as affect. Children with motor speech disorders display particular deficits in prosodic aspects of speech. Shriberg and his colleagues (Shriberg, Aram, and Kwiatkowski, 1997) found that a significant proportion of children diagnosed with DAS demonstrated errors characterized by even or misplaced stress in their spontaneous speech. In a study using direct measures of lip and jaw movement during the production of different stress patterns, Goffman (1999) reported that children with a diagnosis of SLI, who also demonstrated speech production and morphological errors, were poor at producing large and small movements sequentially across different stress contexts. For example, in the problematic weak-strong prosodic sequence, these children had difficulty producing small movements corresponding to unstressed syllables. Overall, the control of movement for the production of stress is a frequently cited deficit in children with motor speech disorders.

**General Motor Development.** In the clinical literature, general neuromotor status has long been implicated as contributing to even relatively subtle speech and language deficits (Morris and Klein, 1987). Empirical studies have provided evidence that aspects of gross and fine motor (e.g., peg moving, gesture imitation) performance are below expected levels in children with variable speech errors, DAS (Bradford and Dodd, 1996), and many diagnosed with SLI (Bishop and Edmundson, 1987; Hill, 2001). Such findings suggest that many speech production disorders include a general motor component.

As is apparent, an understanding of speech motor contributions to the acquisition of speech and language is in its infancy. However, it is clear that intervention approaches for these children need to incorporate motor as well as language components. Although efficacy studies are scarce, several investigators have proposed techniques for the treatment of motor speech disorders in children. Although the emphasis has been on DAS, these approaches could be tailored to more general motor speech deficits. Major approaches to intervention have focused on motor programming (Hall, Jordan, and Robin, 1993) and tactile-kinesthetic and rhythmic (Square, 1994) deficits. Hierarchical language organization has also been emphasized, supporting the intimate
links between linguistic and movement variables (Velleman and Strand, 1994).

New models of speech and language development are needed that integrate motor and language variables in a way that is consistent with recent neurophysiological and behavioral evidence. Further, new methods of recording respiratory, laryngeal, and articulatory behaviors of infants and young children during the production of meaningful linguistic activity should provide crucial data for understanding how language and motor components of development interact across normal and disordered development. Such tools should also help answer questions about appropriate interventions for children whose deficits are influenced by atypical motor control processes.

See also DEVELOPMENTAL APRAXIA OF SPEECH.

—Lisa Goffman

References


Further Readings


Mutism, Neurogenic

Mutism is speechlessness. It can be neurologic or behavioral. Neurogenic mutism is a sign and can result from many developmental or acquired nervous system diseases and conditions. It usually accompanies other signs, but in rare cases it appears in isolation. All levels of the neuroaxis from the brainstem to the cortex have been implicated. Damage to any of the putative processes critical to speech, including intention, motor programming, and execution, as well as linguistic and prelinguistic processes have been invoked to explain mutism’s appearance. A reasonably traditional review of the syndromes and conditions of which mutism is a frequent part and traditional and emerging explanations for mutism’s appearance are offered here.

Definition. Duffy defines mutism traditionally as “the absence of speech” (1995, p. 282). Von Cramon (1981) adds the inability to produce nonverbal utterances. Lebrun (1990) requires normal or relatively preserved comprehension. Gelabert-Gonzalez and Fernandez-Villa (2001) require “unimpaired consciousness” (p. 111). Each of these definitions has strengths, but none dominates the literature. Therefore, the literature can be a bit of a muddle. The literature is also challenging because of the mutism population’s heterogeneity. One response to this heterogeneity has been to classify mutism according to relatively homogeneous subtypes.

Traditionally, groupings of mute patients have been organized according to the putative pathophysiology (Turkstra and Bayles, 1992), by syndrome, by etiology, or by a mix of syndrome and medical etiology (Lebrun, 1990; Duffy, 1995). This last approach guides the organization of the following discussion.

Akinetic Mutism. Akinetic mutism (AM) is a syndrome of speechlessness and general akinesia that exists in the context of residual sensory, motor, and at least some cognitive integrity and a normal level of arousal. The designation abulic state may be a synonym (Duffy, 1995), as are apalllic state and coma vigil. Persons with AM often are silent, despite pain or threat. Bilateral and occasionally unilateral left or right anterior cerebral artery occlusion with involvement of the anterior cingulate gyrus or supplementary motor area is frequently implicated (Nicolai, van Putten, and Tavy, 2001). Recent data suggest that the critical areas are the portions of the medial frontal lobes immediately anterior to the supplementary motor area and portions of the anterior cingulate gyrus above the most anterior body of the corpus callosum. These regions appear to be involved in gating intention (plans of action) (Picard and Strick, 1996; Cohen et al., 1999). Lesions of the globus pallidus, thalamus, and other subcortical structures can also result in AM. Schiff and Plum (2000) advocate for a companion syndrome of “hyperkinetic mutism” resulting most frequently from bilateral temporal, parietal, and occipital junction involvement in which the patient is speechless but moving. The cause may be any nervous system–altering condition, including degenerative diseases such as Creutzfeldt-Jacob disease (Otto et al., 1998), that alters what Schiff and Plum (2000) posit to be a series of corticostriatopallidal-thalamocortical (CSPTC) loops. CSPTC loops are critical to triggering or initiating vocalization (Mega and Cohenour, 1997) and to the drive or will to speak. AM is to be differentiated from persistent vegetative state, which reflects extensive damage to all cerebral structures, most critically the thalamus, with preservation of brainstem function (Kinney et al., 1994).

Mutism in Aphasia. Mutism can be a feature of severe global aphasia. Patients with severe anomia, most often in relation to thalamic lesions, may initially exhibit no capacity for spontaneous language and little for naming, but they can repeat. Certain types of transcortical motor aphasia (Alexander, Benson, and Stuss, 1989), most particularly the adynamic aphasia of Luria (1970), may be associated with complete absence of spontaneous speech. However, these patients are reasonably fluent during picture description, and this syndrome likely reflects a prelinguistic disorder involving defective spontaneous engagement of concept representations (Gold et al., 1997).

Mutism in Apraxia. Immediately after stroke, a profound apraxia of speech (called aphemia and by a variety of other names in the world’s literature) can cause mutism, as can primary progressive apraxia of speech. In the acute stage of stroke, the mutism is thought to signal an apraxia of phonation. The hypothesis is that mutism due to apraxia reflects a profound failure of motor programming.

Mutism in Dysarthria. Mutism can be the final stage of dysarthria (anarthria) in degenerative diseases such as amyotrophic lateral sclerosis, slowly progressive anarthria (Broussolle et al., 1996), olivopontocerebellar atrophy, Parkinson’s disease, Shy-Drager syndrome, striatonigral degeneration, and progressive supranuclear palsy (Nath et al., 2001). Speech movements are impossible because of upper and lower motor neuron
destruction. Cognitive changes may hasten the mutism in some of these degenerative diseases. Anarthric mutism can be present at onset and chronically in locked-in syndrome (Plum and Posner, 1966) and the syndrome of bilateral infarction of opercular motor cortex. Relatively recently, a syndrome beginning with mutism and evolving to dysarthria, called temporary mutism followed by dysarthria (TMFD) (Orefice et al., 1999) or mutism and subsequent dysarthria syndrome (MSD) (Dunwoody, Alsagoff, and Yuan, 1997), has been described. Pontomesencephalic stroke is one cause.

**Mutism in Dementia.** Mutism has been reported in Alzheimer’s disease, cerebrovascular dementia, and most frequently and perhaps earliest in frontotemporal dementia (Bathgate et al., 2001). It can also occur in other cortico-subcortical degenerative diseases, including corticobasal degeneration. Its occurrence in the late stages of these conditions is predictable, based on the hierarchical organization of cognitive, linguistic, and speech processes. When cognitive processes are absent or severely degraded, speech does not occur.

**Mutism Post Surgery.** Mutism can occur after neurosurgery (Pollack, 1997; Siffert et al., 2000). So-called cerebellar mutism can result from posterior fossa surgery, for example. It is hypothesized that disruption of connections between the cerebellum, thalamus, and supplementary motor area causes impaired triggering of vocalization (Gelabert-Gonzalez and Fernandez-Villa, 2001). Mutism may also occur after callosotomy (Sussman et al., 1983), perhaps because of damage to frontal lobe structures and the cingulate gyrus.

**Mutism in Traumatic Brain Injury.** Mutism is frequent in traumatic brain injury. Von Cramon (1981) called it the “traumatic midbrain syndrome.” He speculated that the mechanism is “temporary inhibition of neural activity within the brain stem vocalization center” (p. 804) within the pontomesencephalic area. Often mutism is followed by a period of whispered speech in this population.

**Summary.** Speech depends on myriad general and specific cognitive, motor, and linguistic processes. These processes are widely distributed in the nervous system. However, the frontal lobes and their connections to subcortical and brainstem structures are the most critical. Mutism, therefore, is common, but not inevitable, as an early, late, or chronic sign of damage, regardless of type, to these mechanisms.

—John C. Rosenbek

**References**


**Further Readings**

Bannur, U., and Rajshekhar, V. (2000). Post operative supple-


Van der Nerwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sen-

**Orofacial Myofunctional Disorders in Children**

Orofacial myology is the scientific and clinical knowledge related to the structure and function of the muscles of the mouth and face (orofacial muscles) (American Speech-Language-Hearing Association [ASHA], 1993). Orofacial myofunctional disorders are characterized by abnormal fronting of the tongue during speech or swallowing, or when the tongue is at rest. ASHA defines an orofacial myofunctional disorder as “any pattern involving oral and/or orofacial musculature that interferes with normal growth, development, or function of structures, or calls attention to itself” (ASHA, 1993, p. 22). With orofacial myofunctional disorders, the tongue moves forward in an exaggerated way and may protrude from the upper and lower teeth during speech, swallowing, or at rest. This exaggerated tongue fronting is also called a tongue thrust or a tongue thrust swallow and may contribute to malocclusion, lisping, or both (Young and Vogel, 1983; ASHA, 1989).

A tongue thrust type of swallow is normal for infants. The forward tongue posture typically diminishes as the child grows and matures. Orofacial myofunctional disorders may also be due to lip incompetence, which is a “lips-apart resting posture or the inability to achieve a lips-together resting posture without muscle strain” (ASHA, 1993, p. 22). During normal development, the lips are slightly separated in children. With orofacial myofunctional disorders, a lips-apart posture persists.

Orofacial myofunctional disorders may be due to a familial genetic pattern that determines the size of the mouth, the arrangement and number of teeth, and the strength of the lip, tongue, mouth, or face muscles (Hanson and Barrett, 1988). Environmental factors such as allergies may also lead to orofacial myofunctional disorders. For example, an open mouth posture may result from blocked nasal airways due to allergies or enlarged tonsils and adenoids. The open-mouth breathing pattern may persist even after medical treatment for the blocked airway. Other environmental causes of orofacial myofunctional disorders may be excessive thumb or finger sucking, excessive lip licking, teeth clenching, and grinding (Van Norman, 1997; Romero, Bravo, and Perez, 1998). Thumb sucking, for example, may change the shape of a child’s upper and lower jaw and teeth, requiring speech, dental, and orthodontic intervention (Umberger and Van Reenen, 1995; Van Norman, 1997).
The severity of the problem depends on how long the habit is maintained.

Typically, a team of professionals, including a dentist, orthodontist, physician, and speech-language pathologist, is involved in the assessment and treatment of children with orofacial myofunctional disorders (Benkert, 1997; Green and Green, 1999; Paul-Brown and Clausen, 1999). Assessment is conducted to diagnose normal and abnormal parameters of oral myofunctional patterns (ASHA, 1997). The dentist focuses on the effect of pressure of the tongue against the gums; this kind of tongue pressure may interfere with the normal process of tooth eruption. An orthodontist may be involved when the tongue pressure interferes with alignment of the teeth and jaw. A physician needs to verify that an airway obstruction is not causing the tongue thrust. Speech-language pathologists assess and treat swallowing disorders, speech disorders, or lip incompetence that result from orofacial myofunctional disorders. As with all other assessment and treatment processes, speech-language pathologists need to have the appropriate training, education, and experience to practice in the area of orofacial myofunctional disorders (ASHA, 2002).

An orofacial myofunctional assessment is typically prompted by referral or a failed speech screening for a child older than 4 years of age. Assessment should be based on orofacial myofunctional abilities and education, vocation, social, emotional, health, and medical status. An orofacial myofunctional assessment by a speech-language pathologist typically includes the following procedures (ASHA, 1997, p. 54):

- Case history
- Review of medical/clinical health history and status (including any structural or neurological abnormalities)
- Observation of orofacial myofunctional patterns
- Instrumental diagnostic procedures
- Structural assessment, including observation of the face, jaw, lips, tongue, teeth, hard palate, soft palate, and pharynx
- Perceptual and instrumental measures to assess oral and nasal airway functions as they pertain to orofacial myofunctional patterns and/or speech production (e.g., speech articulation testing, aerodynamic measures)

Speech may be unaffected by orofacial myofunctional disorders (Khinda and Grewal, 1999). However, some speech sound errors, called speech misarticulations, may be causally related to orofacial myofunctional disorders. The sounds most commonly affected by orofacial myofunctional disorders include s, z, sh, zh, ch, and j. Sound substitutions (e.g., th for s, as in “thun” for “sun”) or sound distortions may occur. A weak tongue tip may result in difficulties producing the sounds t, d, n, and l.

Speech-language pathologists evaluate speech sound errors resulting from orofacial myofunctional disorders, as well as lip incompetence and swallowing disorders (ASHA, 1991). The assessment information is used to develop appropriate treatment plans for individuals who are identified with orofacial myofunctional disorders. Before speech and swallowing treatment is initiated, medical treatment may be necessary if the airway is blocked due to enlarged tonsils and adenoids or allergies. Excessive and persistent oral habits, such as thumb and finger sucking or lip biting, may also need to be eliminated or reduced before speech and swallowing treatments are initiated.

Some speech and swallowing treatment techniques include

- Increasing awareness of mouth and facial muscles.
- Increasing awareness of mouth and tongue postures.
- Completing an individualized oral muscle exercise program to improve muscle strength and coordination. Treatment strategies may include alternating of tongue and lip resting postures and muscle retraining exercises (ASHA, 1997, p. 69).
- Establishing normal speech articulation.
- Establishing normal swallowing patterns. Treatment strategies may include modification of handling and swallowing of solids, liquids, and saliva (ASHA, 1997, p. 69).

The expected outcome of treatment is to improve or correct the patient’s orofacial myofunctional swallowing and speech patterns. Orofacial myofunctional treatment may be conducted concurrently with speech treatment.

Orofacial myofunctional treatment is effective in modifying tongue and lip posture and movement and in improving dental occlusion and a dental open bite or overbite (Christensen and Hanson, 1981; ASHA, 1991; Benkert, 1997). Lip exercises may be successful in treating an open-mouth posture (ASHA, 1989; Pedrazzi, 1997). A combination treatment approach, with a focus on speech correction as well as exercises to treat tongue posture and swallowing patterns, appears to be the optimal way to improve speech and tongue thrust (Umberger and Johnston, 1997). The length of treatment varies according to the severity of the disorder, the age and maturity of the patient, and the timing of treatment in relation to orthodontia. Typically 14–20 sessions or more may occur over a period of 3 months to a year (ASHA, 1989). The value of early treatment is emphasized in the literature (Pedrazzi, 1997; Van Norman, 1997).

ASHA has identified the basic content areas to be covered in university curricula to promote competency in the assessment and treatment of orofacial myofunctional disorders (ASHA, 1989, p. 92), including the following:

1. Oral-facial-pharyngeal structure, development, and function
2. Interrelationships among oral-vegetative functions and adaptations, speech, and dental occlusion, using interdisciplinary approaches
3. Nature of atypical oral-facial patterns and their relationship to speech, dentition, airway competency, and facial appearance
4. Relevant theories such as those involving oral-motor control and dental malocclusion
5. Rationale and procedures for assessment of oral myofunctional patterns, and observation and participation in the evaluation and treatment of patients with orofacial myofunctional disorders
6. Application of current instrumental technologies to document clinical processes and phenomena associated with orofacial myofunctional disorders
7. Treatment options

A Joint Committee of ASHA and the International Association of Orofacial Myology has also delineated the knowledge and skills needed to evaluate and treat persons with orofacial myofunctional disorders (ASHA, 1993). The tasks required include the following:

- Understanding dentofacial patterns and applied physiology pertinent to orofacial myology
- Understanding factors causing, contributing, or related to orofacial myology
- Understanding basic orthodontic concepts
- Understanding interrelationships between speech and orofacial myofunctional disorders
- Demonstrating competence in comprehensive assessment procedures and in identifying factors affecting prognosis
- Demonstrating competence in selecting an appropriate, individualized, criterion-based treatment plan
- Demonstrating a clinical environment appropriate to the provision of professional services
- Demonstrating appropriate documentation of all clinical services
- Demonstrating professional conduct within the scope of practice for speech-language pathology (ASHA, 2001)

Further information on oral myofunction and oral myofunctional disorders is available from ASHA's Special Interest Division on Speech Science and Orofacial Disorders (www.asha.org) and the International Association of Orofacial Myology (www.iaom.com).

—Diane Paul-Brown

References


Further Readings


Phonetic transcription entails using special symbols to create a precise written record of an individual’s speech. The symbols that are most commonly used are those of the International Phonetic Alphabet (IPA), first developed in the 1880s by European phoneticians. Their goal was to provide a different symbol for each unique sound, that is, to achieve a one-to-one correspondence between sound and symbol. For example, because [s] and [ʃ] are phonemically distinct in some languages, such as English, they are represented differently in the phonetic alphabet. Thus, the elongated s is used for the voiceless palatoalveolar fricative, as in [ʃ], to differentiate it from the voiceless alveolar fricative, as in [s].

The IPA has undergone several revisions since its inception but remains essentially unchanged. In the familiar consonant chart, symbols for pulmonic consonants are organized according to place of articulation, manner of articulation, and voicing. Nonpulmonic consonants, such as clicks and ejectives, are listed separately, as are vowels, which are shown in a typical vowel quadrangle. Symbols for suprasegmentals, such as length and tone, are also provided, as are numerous diacritics, such as [ə] for a dentoalveolar [s].

The most recent version of the complete IPA chart can be found in the Handbook of the International Phonetic Association (IPA, 1999) as well as in a number of phonetics books (e.g., Ladefoged, 2001; Small, 1999). Illustrations of the sounds of the IPA are available through various sources, such as Ladefoged (2001) and Wells and House (1995). In addition, training materials and phonetic fonts can be downloaded from the Internet. Some new computers now come equipped with “Unicode” phonetic symbols.

Although extensive, the IPA does not capture all of the variations that have been observed in children’s speech. For this reason, some child/clinical phonologists have proposed additional symbols and diacritics (e.g., Bush et al., 1973; Edwards, 1986; Shriberg and Kent, 2003). The extended IPA (extIPA) was adopted by the International Clinical Phonetics and Linguistics Association (ICPLA) Executive Committee in 1994 to assist in and standardize the transcription of atypical speech (e.g., Duckworth et al., 1990). The extIPA includes symbols for sounds that do not occur in “natural” languages, such as labiodental and interdental plosives, as well as many diacritics, such as for denasalized and unaspirated sounds. It also includes symbols for transcribing connected speech (e.g., sequences of quiet speech, fast or slow speech), as well as ways to mark features such as silent articulation. Descriptions and examples can be found in Ball, Rahilly, and Tench (1996) and Powell (2001).

When transcribing child or disordered speech, it is sometimes impossible to identify the exact nature of a segment. In such cases, “cover symbols” may be used. These symbols consist of capital letters to represent major sound classes, modified with appropriate diacritics. Thus, an unidentifiable voiceless fricative can be transcribed with a capital F and a small under-ring for voicelessness (e.g., Stoel-Gammon, 2001).

Relatively little attention has been paid to the transcription of vowels in children’s speech (see, however, Pollock and Berni, 2001). Even less attention has been paid to the transcription of suprasegmentals or prosodic features. Examples of relevant IPA and extIPA symbols appear in Powell (2001), and Snow (2001) illustrates special symbols for intonation.
Broad or “phonemic” transcriptions, which capture only the basic segments, are customarily written in slashes (virgules), as in /paɪ/ or /teləfon/. “Narrow” or “close” transcriptions, which often include diacritics, are written in square brackets. A narrow transcription more accurately represents actual pronunciation, whether correct or incorrect, as in [pʰæɪ] for pie, with aspiration on the initial voiceless stop, or a young child’s rendition of star as [tʰɔː] or fish as [fiːs].

How narrow a transcription needs to be in any given situation depends on factors such as the purpose of the transcription, the skill of the transcriber, and the amount of time available. As Powell (2001) points out, basic IPA symbols are sufficient for some clinical purposes, for example, if a client’s consonant repertoire is a subset of the standard inventory. A broad transcription is generally adequate to capture error patterns that involve deletion, such as final consonant deletion or cluster reduction, as well as those that involve substitutions of one sound class for another, such as gliding of liquids or stopping of fricatives.

If no detail is included in a transcription, however, the analyst may miss potentially important aspects of the production. For instance, if a child fails to aspirate initial voiceless stops, the unaspirated stops should be transcribed with the appropriate (extIPA) diacritic (e.g., [pː], [tː], as in [pːi] for pea). Such stops can easily be mistaken for the corresponding voiced stops and erroneously transcribed as [b], [d], and so on. The clinician might then decide to work on initial voicing, using minimal pairs such as pea and bee. This could be frustrating for a child who is already making a subtle (but incorrect) contrast, for example, between [pː] and [b].

To give another example, a child who is deleting final consonants may retain some features of the deleted consonants as “marking” on the preceding vowel, for instance, vowel lengthening (if voiced obstructions are deleted) or nasalization (if nasal consonants are deleted). Unless the vowels are transcribed narrowly, the analyst may miss important distinctions, such as between [bi] (beet), [biː] (bead), and [bi] (bean).

Stoel-Gammon (2001) suggests using diacritics only when they provide additional information, not when they represent adult-like use of sounds. For example, if a vowel is nasalized preceding a nasal consonant, the nasalization would not need to be transcribed. However, if a vowel is nasalized in the absence of a nasal consonant, as in the preceding example, or if inappropriate nasalization is observed, a narrow transcription is crucial.

Phonetic transcription became increasingly important for speech-language pathologists with the widespread acceptance of phonological assessment procedures in the 1980s and 1990s. Traditional articulation tests (e.g., Goldman and Fristoe, 1969) did not require much transcription. Errors were classified as substitutions, omissions, or distortions, and only the substitutions were transcribed. Therefore, no narrow transcription was involved.

In order to describe patterns in children’s speech, it is necessary to transcribe their errors. Moreover, most phonological assessment procedures require whole word transcription (e.g., Hodson, 1980; Khan and Lewis, 1986), so that phonological processes involving more than one segment, such as assimilation (as in [gæk] for truck), can be more easily discerned. (In fact, Shriberg and Kwiatkowski, 1980, use continuous speech samples, necessitating transcription of entire utterances.)

To facilitate whole word transcription, some clinical phonologists, such as Hodson (1980) and Louko and Edwards (2001), recommend writing out broad transcriptions of target words (e.g., /træk/) ahead of time and modifying them “on line” for a tentative live transcription that can be verified or refined by reviewing a tape of the session. Although this makes the transcription process more efficient, it can also lead the transcriber to mishear sounds or to “hear” sounds that are not there (Oller and Eilers, 1975). Louko and Edwards (2001) provide suggestions for counteracting the negative effects of such expectation.

If a speech-language pathologist is going to expend the time and energy necessary to complete a phonological analysis that is maximally useful, the transcription on which it is based must be as accurate and reliable as possible. Ideally, the testing session should be audio- or video-recorded on high-quality tapes and using the best equipment available, and it should take place in a quiet environment, free of distractions (see Stoel-Gammon, 2001). Because some sounds are difficult to transcribe accurately from an audiotape (e.g., unreleased final stops), it is advisable to do some transcribing on-line.

One way to enhance the accuracy of a transcription is to transcribe with a partner or to find a colleague who is willing to provide input on difficult items. “Transcription by consensus” (Shriberg, Kwiatkowski, and Hoffman, 1984), although impractical in some settings, is an excellent way to derive a transcription and to sharpen one’s skills. This involves two or more people transcribing a sample at the same time, working independently, then listening together to resolve disagreements.

Sometimes it is desirable to assess the reliability of a transcription. For interjudge reliability, the transcriber relistsens to a portion of the sample at some later time and compares the two transcriptions on a sound-by-sound basis, determining a percent of “point-to-point” agreement. The same procedure may be used for determining interjudge reliability, except that a second listener’s judgments are compared with those of the first transcriber. Reliability rates for children’s speech vary greatly, depending on factors such as the type of sample (connected speech or single words) and how narrow the transcription is, with reliability rates being higher for broad transcription (see Cucchiarini, 1996; Shriberg and Lof, 1991). Alternative methods of assessing transcription agreement may sometimes be appropriate. For instance, in assessing the phonetic inventories of young children, Stoel-Gammon (2001) suggests measuring agreement of features (place or manner) rather than identity of segments.

People who spend long hours transcribing children’s speech often look forward to the day when accurate
computer transcription will become a reality. Although computer programs may be developed to make transcription more objective and time-efficient, speech-language pathologists will continue to engage in the transcription process because of what can be learned through carefully listening to and trying to capture the subtleties of a person’s speech. Therefore, phonetic transcription is likely to remain an essential skill for anyone engaged in assessing and remediating speech sound disorders.

—Mary Louise Edwards

References


Further Readings


Phonological Awareness Intervention for Children with Expressive Phonological Impairments

Phonological awareness refers to an individual’s awareness of the sound structure of a language. Results from a number of studies indicate that phonological awareness skills are highly correlated with reading success (see Stanovich, 1980) and that phonological awareness can be enhanced by direct instruction (see Blachman et al., 1994). Some scientists prefer using the terms phonological sensitivity or metaphonology rather than phonological awareness. These three terms are generally considered comparable in meaning, except that metaphonology implies that the awareness is at a more conscious level. A fourth term, phonemic awareness, refers only to phonemes, whereas phonological awareness includes syllables and intrasyllabic units (onset and rime). Phonological processing, the most encompassing of these related terms, includes phonological production, verbal working memory, word retrieval, spelling, and writing, as well as phonological awareness. Among the individuals who have been identified most consistently as being “at risk” for failure to develop appropriate phonological awareness skills, and ultimately literacy, are children with expressive phonological impairments (EPIs) (Webster and Plante, 1992).

Relationship Between Expressive Phonological Impairment (EPI) and Phonological Awareness. A growing body of evidence indicates that young children with severe EPI go on to experience problems in literacy. As well, results from another line of research indicate that individuals with reading disabilities evidence more phonological production difficulties (e.g., with multisyllabic words) than their peers with typical reading abilities (Catts, 1986). Bird, Bishop, and Freeman (1995) found that the children who had severe EPI experienced greater difficulty with phonological awareness tasks than their ability-matched peers, even when the tasks did not require a verbal response. Clarke-Klein and Hodson (1995) obtained similar results for spelling. Larivee and Catts (1999), who tested children first in kindergarten and again 1 year later, found that expressive phonology (measured by a multisyllabic word and nonword production task) and phonological awareness scores in kindergarten accounted for significant amounts of variance in first-grade reading.

Several investigators (e.g., Bishop and Adams, 1990; Catts, 1993), however, have reported that phonological impairments alone do not have as great an impact on literacy as language impairments do. A possible explanation for this discrepancy may be the level of EPI severity in the participants in their studies.

Severity Considerations. A common practice in the articulation/phonology literature is to report the number of errors on an articulation test. Not all speech sound errors are equal, however. For example, if two children have 16 errors on the same test, some examiners might view them as equal. If, however, one child evidences a lisp for all sibilants and the other has 16 omissions, the impact on intelligibility will be vastly different. Moreover, the child with extensive omissions might be identified as having a language impairment because of the omission of final consonants (which would affect the production of word-final morphemes on an expressive language measure). Some highly unintelligible children who are considered to have a language impairment may, in fact, have a severe phonological impairment with intact receptive language abilities. Typically such children produce final morphemes as they learn the phonological pattern of word-final consonants.

Phonological Awareness Treatment Studies for Children with EPI. Although there have been numerous studies reporting the results of phonological awareness treatment, only a few investigators have focused on children with phonological or language impairments. van Kleeck, Gillam, and McFadden (1998) provided classroom-based phonological awareness treatment (15 minutes twice a week) to 16 children with speech and/or language disorders (8 in a preschool class and 8 in a pre-kindergarten class). The small-group sessions focused on rhyming during the first semester and on phoneme
awareness during the second semester. The treatment groups and a nontreatment comparison group all made substantial gains in rhyming. Children in the treatment groups, however, made markedly greater gains on phonemic awareness tasks than children in the nontreatment group. Information on changes in expressive phonology or language was not provided by the investigators.

Howell and Dean (1994) used their Metaphon program to provide both phonological awareness and production treatment for 13 preschool children with EPI in Scotland. In phase 1 of this program, children progress from the concept/sound (not speech) level to the phoneme level to the word level. Minimal pairs are used extensively during phase 1. In phase 2, the progression is from word level to sentence level. The children attended between 11 and 34 30-minute sessions weekly. Single subject case study results indicated that the children improved on both phonological production and phonological awareness tasks (sentence and phoneme segmentation).

Harbers, Paden, and Halle (1999) provided individual treatment to four preschool children with EPI for 6–9 months that focused on both feature awareness and production for three phonological patterns that the children lacked. All four children targeted /s/ clusters. Three targeted strident singletons, two targeted velars, two targeted liquids, and one targeted final consonants. The investigators used a combination of the Metaphon (Howell and Dean, 1994) and Cycles (Hodson and Paden, 1991) treatment approaches. Improvement in the production of /s/ clusters coincided with gains in recognizing /s/ cluster features for two of the four children targeting /s/ clusters. Both of the children targeting velars also evidenced concomitant gains in production and awareness. For the remaining targets, there was a slight tendency for the two variables (phonological awareness and production) to move in similar directions, but inconsistencies occurred.

Gillon (2000) conducted a phonological awareness treatment study in New Zealand that involved 91 children with “spoken language impairment” between the ages of 5 and 7 years. Twenty-three children participated in an experimental “integrated” treatment program. A second group of 23 children received traditional speech–sound treatment. Two additional groups served as controls. One treatment group of 15 children who received “minimal” intervention, and the other consisted of 30 phonologically normal children. Children in the first treatment group received two 60-minute sessions per week until a total of 20 hours of intervention had been completed. The second group participated in phoneme-oriented sessions for the same amount of time. All of the children continued participating in their regular classroom literacy instruction, which was based on a “Whole Language” model.

The children in the first group did not receive direct production treatment for EPI during the course of the study. Additional stimulus items for children’s individual speech sound errors were integrated into some of the activities, however. The phonological awareness treatment focused on the development of skills at the phonemic level and integrated phonological awareness activities with grapheme-phoneme correspondence training. Activities included (a) picture Bingo and oddity games for rhyme awareness, (b) identification of initial and final sounds, and sometimes medial sounds, (c) phoneme segmentation, (d) phoneme blending, and (e) linking speech to print. The children in this group made significantly greater gains in phonological awareness and reading scores than the children in the other groups. Moreover, the children also made greater gains in phonological production than children in the other groups with EPI. The results of this investigation lend support to the contention that it is important to incorporate phonological awareness tasks into treatment sessions for children with EPI.

Enhancing Phonological Awareness Skills. Available tasks range in difficulty from simple “yes-no” judgments regarding whether two words rhyme to complex phonological manipulation activities (e.g., pig Latin, spoonerisms). Moreover, many activities that are commonly used in treatment sessions have phonological awareness components. When children are taught how a sound is produced and how it feels, they develop awareness about place, manner, and voicing aspects of the sounds in their phonological system. One phonological awareness treatment program (Lindamood and Lindamood, 1998) has a component that specifically addresses teaching the articulatory characteristics of phonemes to all children with reading disabilities, even when there are no phonological production problems. Moreover, when children learn about where a sound is located in a word (initial, medial, or final position), they develop awareness about word positions.

One phonological awareness activity that has proved to be particularly effective is the “Say-It-And-Move-It” task, using Elkonin cards (Ball and Blachman, 1991, adapted from Elkonin, 1963). Children are taught to represent the sounds in one- (e.g., /a/), two- (e.g., /up/), or three-phoneme (e.g., /cat/) words by using manipulatives. Initially blank tiles or blocks are used. Tiles with graphemes are incorporated after the child demonstrates recognition of the sounds for the letters. The top half of the paper has a picture of a word. The bottom half has the appropriate number of boxes for the phonemes needed for the word. Children are taught to say each word slowly and to move one manipulative for each sound into the boxes from left to right.

Another phonological awareness activity that is widely used both for assessment and for segmentation practice is categorization. This task requires matching and oddity awareness skills. Typically the child is given four pictures and is to identify the one that does not match the others in some aspect (e.g., rhyme) and thus is the “odd one out.” Categorization also is used for individual sounds (e.g., initial consonants).

Learning to blend phonological segments to make words is another important task and one that is extremely difficult for some children. Blending tasks com-
monly start at the word level with compound words (e.g., ice plus cream) followed by blending syllables (e.g., can plus dee; candy). Blending intrasyllabic units (e.g., onset and rime, as sh plus eep; and body and coda, as shee plus p) should precede blending individual phonemes (e.g., sh plus ee plus p).

Another task that has been found to be highly correlated with success in reading is deletion (e.g., elision task, Rosner and Simon, 1971). As with blending, it is important to begin with the larger segments (e.g., compound words). The child says the word (e.g., cowboy), and then, after part of the word is removed (e.g., boy), says the new word (cow). After a child demonstrates success at the larger unit levels, individual phonemes are deleted (e.g., take away /t/ from note), leaving no

The task that consistently has accounted for the greatest amount of variance in predicting decoding success is manipulation. Children who are most successful performing phoneme manipulation tasks such as spoonerisms typically are the best decoders (Strattman, 2001). Phonological manipulation in “pattern” songs (e.g., “Apples and Bananas”) seems to be an extremely enjoyable task for very young children and can help them be more aware of sounds and word structures.

**Implications for Best Practices.** Because children with EPI appear to be at risk for the development of normal reading and writing skills even after they no longer have intelligibility issues, it seems prudent to incorporate activities to enhance phonological awareness skills while they are receiving treatment for phonological production. Moreover, results from Gillon’s (2000) study indicate that enhancing phonological awareness skills leads to improvement in phonological production. Thus, enhancing phonological awareness skills appears to serve a dual purpose for children with expressive phonological impairments.

—Barbara Hodson and Kathy Strattman

**References**


**Further Readings**


**Phonological Errors, Residual**

Shriberg (1994) has conceptualized developmental phonological disorders as speech disorders that originate during the developmental period. In most cases the cause of such disorders cannot be attributed to significant involvement of a child’s speech or hearing processes, cognitive-linguistic functions, or psychosocial processes (Bernthal and Bankson, 1998), but causal origins may be related to genetic or environmental differences (Shriberg, 1994; Shriberg and Kwiatkowski, 1994). Children with developmental phonological disorders are heterogeneous and exhibit a range in the severity of their phonological disorders. Generally, the expected developmental period for speech sound acquisition ends at approximately 9 years of age, thus encompassing birth through the early school years. In sum, it is posited that children who exhibit phonological disorders differ with regard to the etiology and severity of the disorder and include both preschool and school-age children (Deputy and Weston, 1998). Some individuals with developmental phonological disorders acquire normal speech, while others continue to exhibit a phonological disorder throughout the life span, despite having received treatment for the phonological disorder (Shriberg et al., 1997).

Residual phonological errors are a subtype of developmental phonological disorders that persist beyond the expected period of speech-sound development or normalization (Shriberg, 1997). They are present in the speech of older school-age children and adults. Individuals with residual errors can be further classified into subgroups of those with a history of speech delay and those without a history of speech delay (i.e., individuals in whom a speech delay was diagnosed at some time during the developmental period and those who were not so diagnosed). It is postulated that the two groups differ with respect to causal factors. The residual errors of the first group are thought to reflect environmental influences, while nonenvironmental causal factors such as genetic transmission are thought to be responsible for the phonological errors of the second group.

Most residual errors have been identified as distortions (Smit et al., 1990; Shriberg, 1993) of the expected allophones of a particular phoneme. Distortions are variant productions that do not fall within the perceptual boundaries of a specific target phoneme (Danillof, Wilcox, and Stephens, 1980; Bernthal and Bankson, 1998). It has been hypothesized that distortions reflect incorrect allophonic rules or sensorimotor processing limitations. That is, such productions are either permanent or temporary manifestations of inappropriate allophonic representation and/or the sensorimotor control of articulatory accuracy. It has been suggested that children initially delete and substitute sounds and then produce distortions of sounds such as /r/, /l/, and /s/ when normalizing sound production; however, investigative study has not supported this hypothesis as a generality in children who normalize their phonological skills with treatment (Shriberg and Kwiatkowski, 1988). Ohde and Sharf (1992) provide excellent descriptions of the acoustic and physiologic parameters of common distortion errors.

Smit et al. (1990) conducted a large-scale investigation of speech sound acquisition and reported that the distortion errors noted in the speech of their older test subjects varied with respect to judged clinical impact or severity. Some productions were judged to be minor distortions, while others were designated as clinically significant. Shriberg (1993) also noted such differences in his study of children with developmental phonological disorders. He classified the errors into nonclinical and clinical distortion types. Nonclinical distortions are thought to reflect dialect or other factors such as speech-motor constraints and are not targeted for therapy. Clinical distortions are potential targets for treatment and have been categorized by prevalence into common and uncommon types. The most common and uncommon types are listed in Figure 1. The most common residual errors include distortions of the sound classes of liquids, fricatives, and affricates. Uncommon distortion errors include errors such as weak or imprecise consonant production and difficulty maintaining nasal and voicing features. In most cases, residual errors constitute...
Common Distortion Errors
1. Dentilation of voiced/voiceless sibilant fricatives or affricates
2. Derhotacized /r/, /ʃ/, /ʒ/
3. Lateralization of voiced/voiceless sibilant fricatives or affricates
4. Velarized /l/ or /r/
5. Labialized /l/ or /r/

Uncommon Distortion Errors
1. Weak consonant productions
2. Imprecise articulation of consonants and vowels
3. Inability to maintain oral/nasal contrasts
4. Difficulty in maintaining correct voicing contrasts

Figure 1. Common and uncommon distortion errors as reported by Shriberg (1993).

minor involvement of phonological production and do not have a significant impact on intelligibility, but research indicates that normal speakers react negatively to persons with even minor residual errors (Mowrer, Wahl, and Doolan, 1978; Silverman and Paulus, 1989; Crowe Hall, 1991).

Treatment for persons with residual errors is generally carried out using approaches that have been used with younger children. The treatment approaches are based on motor learning or cognitive-linguistic concepts (Lowe, 1994; Bauman-Waengler, 2000); however, in most cases a motor learning approach is utilized (Gierut, 1998). Although most individuals normalize their residual errors with intervention, some individuals do not (Dagenais, 1995; Shuster, Ruscello, and Toth, 1995). The actual number of clients in the respective categories is unknown, but survey data of school practitioners reported by Ruscello (1995a) indicate that a subgroup of clients do not improve with traditional treatment methods. Respondents indicated that children either were unable to achieve correct production of an error sound or achieved correct production but were unable to incorporate the sound into spontaneous speech. The respondents did not list the types of sound errors, but the error sounds reported are in agreement with the residual errors identified by both Shriberg (1993) and Smit et al. (1990).

In some cases, specially designed treatments are necessary to facilitate remediation of residual errors. For example, principles from biofeedback and speech physiology have been incorporated into treatments (Dagenais, 1995; Ruscello, 1995b; Gibbon et al., 1999). Different forms of sensory information other than auditory input have been provided to assist the individual in developing appropriate target productions. Shuster, Ruscello, and Toth (1995) identified two older children with residual /ʃ/ errors who had received traditional long-term phonological treatment without success. A biofeedback treatment utilizing real-time spectrography was implemented for both subjects, and the results indicated that the two subjects were able to acquire correct production of the former residual error.

In summary, residual errors are a distinct subtype of developmental phonological errors that are present in the speech of older children and adults who are beyond the period of normal sound acquisition. Most residual errors are described as distortions, which are sound variations that are not within the phonetic boundaries of the intended target sound. Generally, residual errors are minor in terms of severity and do not interfere with intelligibility, but normal speakers do react negatively to such minor speech variations. An exact estimate of children and adults with residual errors is unknown, but it is thought that there are substantial numbers of individuals with such a phonological disorder.

—Dennis M. Ruscello

References


**Further Readings**


**Phonology: Clinical Issues in Serving Speakers of African-American Vernacular English**

Word pronunciation is an overt speech characteristic that readily identifies dialect differences among normal speakers even when other aspects of their spoken language do not. Although regional pronunciation differences in the United States were recognized historically, social dialects were not. Nonprestige social dialects in particular were viewed simply as disordered speech. A case in point is the native English dialect spoken by many African Americans, a populous ethnic minority group. This dialect is labeled in various ways but is referred to here as African American Vernacular English (AAVE). As a result of litigation, legislation, and social changes beginning in the 1960s, best clinical practice now requires speech clinicians to regard social dialect differences in defining speech norms for clinical service delivery. This mandate has created challenges for clinical practices.

One clinical issue is how to identify AAVE speakers. African Americans are racially, ethnically, and linguistically diverse. Not all learn AAVE, and among those who do, the density of use varies. This discussion considers only those African Americans with an indigenous slave history in the United States and ancestral ties to Sub-Saharan Africa. The native English spoken today is rooted partly in a pidgin-creole origin. Since slavery was abolished, the continuing physical and social segregation of African Americans has sustained large AAVE communities, particularly in southern states.

Contemporary AAVE pronunciation is both like and unlike Standard English (SE). In both dialects, the vowel and consonant sounds are the same (with a few exceptions), but their use in words differs (Wolfram, 1994; Stockman, 1996b). Word-initial single and clustered consonants in AAVE typically match those in SE except for interdental fricatives (e.g., this > /θ/). The dialects differ in their distributions of word-final consonants. Some final consonants in AAVE are replaced (cf. bath and bathe > /ð/ and /v/, respectively). Others are absent as single sounds (e.g., man) or in consonant clusters (test > /tɛs/). Yet AAVE is not an open-syllable dialect. Final consonants are variably absent in predictable or rule-governed ways. They are more likely to be absent or reduced in clusters when the following word or syllable begins with another consonant rather than a vowel (Wolfram, 1994) or when a consonant is an alveolar as opposed to a labial or velar stop (Stockman, 1991). In multisyllabic words, unstressed syllables (e.g., away > /-we/) in any position may be absent, depending on grammatical and semantic factors (Vaughn-Cooke, 1986). Consonants may also be reordered in some words (e.g., ask > /s/) and multiple words may be merged phonetically (e.g., fixing to > finna; suppose to > sposta) to function as separate words. These broadly predictable AAVE pronunciation patterns differ enough from SE to compromise its intelligibility for unfamiliar listeners. Intelligibility can be decreased further by co-occurring dialect differences in prosodic or nonsegmental (rhythmic and vocal pitch) features (Taron, 1975; Dejarnette and Holland, 1993), coupled with known grammatical, semantic, and pragmatic ones. Consider just the number of grammatical and phonological differences between SE and AAVE in the following example:

**SE:** They are not fixing to ask for the car /<br>**AAVE:** They not finna ask for the car /<br>the ar nat fiksm tu aesk fó tha ka:/
<no-value>det na finə æks fə də ka:/</no-value>
Enough is known about the complex perceptual judgments of speech intelligibility to predict that the more work listeners have to do to figure out what is being said, the more likely is speech to be judged as unclear.

Identifying atypical AAVE speakers can be difficult, especially if known causes of disordered speech—hearing loss, brain damage, and so on—are absent, as is often the case. Clinicians must know a lot about the dialect to defend a diagnosis. But most clinicians (95%) are not African American and have little exposure to AAVE (Campbell and Taylor, 1992). Misdiagnosing normal AAVE speakers as abnormal is encouraged further by the similarity of their typical pronunciation patterns (e.g., final consonant deletion, cluster reduction, and interdental fricative substitutions) to those commonly observed among immature or disordered SE speakers. However, typically developing African-American speakers make fewer errors on standardized articulation tests as they get older (Ratusnik and Koenigsknecht, 1976; Simmons, 1988; Haynes and Moran, 1989). Still, they make more errors than their predominantly white, age-matched peers (Ratusnik and Koenigsknecht, 1976; Seymour and Seymour, 1981; Simmons, 1988; Cole and Taylor, 1990), and they do so beyond the age expected for developmental errors (Haynes and Moran, 1989). Therefore it is unknown whether the overrepresentation of African Americans in clinical caseloads is due to practitioner ignorance, test bias, or an actual higher prevalence of speech disorders as a result of economic poverty and its associated risks for development in all areas.

The accuracy in identifying articulation/phonological disorders improves when test scores are adjusted for dialect differences (Cole and Taylor, 1990), or when the pronunciation patterns for a child and caregiver are compared on the same test words (Terrell, Arensberg, and Rosa, 1992). However, tests of isolated word pronunciation are not entirely useful, even when nonstandard dialect use is not penalized. They typically provide no contexts for sampling AAVE’s variable pronunciation rules, which can cross word boundaries, as in the case of final consonant absence. Although standardized deep tests of articulation (McDonald, 1968) do elicit paired word combinations, they favor the sampling of abutting consonant sequences (e.g., bus fish), which penalize AAVE speakers even more, given their tendency to delete final consonants that precede other consonants as opposed to vowels (Stockman, 1993, 1996b). These issues have encouraged the use of criterion-referenced evaluations of spontaneous speech samples for assessment (see Stockman, 1996a, and Schraeder et al., 1999).

Despite the assessment challenges, it is readily agreed that some AAVE speakers do have genuine phonological/articulatory disorders (Seymour and Seymour, 1981; Taylor and Peters, 1986). They differ from typically developing community peers in both the frequency and patterning of speech sound error. This is true whether the clinical and nonclinical groups are distinguished by the judgments of community informants, such as Head Start teachers (Bleile and Wallach, 1992), other classroom teachers (Washington and Craig, 1992), or speech-language clinicians (Stockman and Settle, 1991; Wilcox, 1996).

AAVE speakers with disorders can differ from their nondisordered peers on speech sounds that are like SE (Type I error, e.g., word-initial single and clustered consonants). They can also differ on sounds that are not like SE either qualitatively (Type II error, e.g., interdental fricative substitutions) or quantitatively (Type III error, e.g., more frequent final consonant absence in abutting consonant sequences). Wolfram (1994) suggested that these three error categories provide a heuristic for scaling the severity of the pronunciation difficulty and selecting targets for treatment.

Two service delivery tracks are within the scope of practice for speech clinicians. One Remediates atypical speech relative to a client’s native dialect. The other one expands the pronunciation patterns of normal speakers who want to speak SE when AAVE is judged to be socially or professionally handicapping (Terrell and Terrell, 1983). For both client populations, effective service delivery requires clinician sensitivity to cultural factors that impact (1) verbal and nonverbal interactions with clients, (2) selection of stimuli (e.g., games and objects) for therapy activities, and (3) scheduling of sessions (Seymour, 1986; Proctor, 1994). However, the service delivery goals do differ for these two populations. For abnormal speakers, the goal is to eradicate and replace existing patterns that decrease intelligible speech in the native dialect. This means that the pronunciation of bath/baθ/ as /baθ/ should not be targeted for change, if it conforms to the client’s target dialect. But a deviation from this expected pronunciation, such as bath/baθ/ > /baθ/ or /βaθ/, is targeted, if observed at an age when developmental errors are not expected. In contrast, the service delivery goal for normal AAVE speakers is to expand rather than eradicate the existing linguistic repertoire (Taylor, 1986). An additive approach assumes that speakers can learn to switch SE and AAVE codes as the communicative situation demands, just as bilingual speakers switch languages. This means that a speaker’s bidialectal repertoire includes both the SE and AAVE pronunciation of “bath” (cf. bath > /baθ/ and /βaθ/).

Meeting these two different service goals requires attention to some issues that are not the same. They affect which patterns are targeted and how change is facilitated. For typical AAVE speakers learning SE, second language acquisition principles are relevant. Besides the production practice, service delivery requires contrastive analysis of the two dialects and attention to sociocultural issues that affect code switching. Correct or target productions are judged relative to SE.

In contrast, for speakers with abnormal pronunciation, AAVE should be targeted. Which features to target in therapy and how to model the input become issues, because most clinicians do not speak AAVE. They also may resist modeling a low social prestige dialect because of negative social attitudes towards it. Wolfram (1994) reminded us that AAVE and SE share many of the same target features (e.g., most word-initial consonants). Errors on shared features (Type I) should be targeted first in treatment. They are likely to impair intelligibility.
even more than the smaller sets of qualitative (Type II) errors, such as stop replacement of interdental fricatives (cf. this /dɪs/ > /bɪs/), or quantitative (Type III) errors, such as final consonant deletion in more than the allowable context number and types. AAVE features should be targeted for treatment only when pronunciation patterns differ from AAVE norms. Articulatory patterns would not be modified if they differed from the clinician’s SE-modeled pattern but matched expected AAVE patterns.

Legitimizing social dialects like AAVE in the United States has required researchers and clinicians to (1) broaden the reference point for normalcy and (2) explore alternative strategies for identifying service needs and modifying word pronunciation. The issues singled out in this entry are not unique to phonological/articulatory problems. However, given their typically higher frequency of occurrence relative to other domains of spoken language in all groups, they may turn up more often in clinical work.

See also DIAPHRAGM SYNDROME: DIALECT VERSUS DISORDER; LANGUAGE DISORDERS IN AFRICAN-AMERICAN CHILDREN.

—Ida J. Stockman

References


Further Readings


Psychosocial Problems Associated with Communicative Disorders

Individuals who study communicative disorders have long been interested in the psychosocial difficulties associated with these problems. This interest has taken different faces over the years as researchers and clinicians have focused on various aspects of the relationship between communicative impairment and psychological and social difficulties. For example, relatively early in the development of the profession of speech-language pathology, some investigators approached specific communicative disorders, such as stuttering, as manifestations of underlying psychological dysfunction (e.g., Travis, 1957). More recent approaches have moved away from considering psychiatric dysfunction as the basis for most speech and language impairment (an exception is alexithymia). Despite this reorientation, there is still considerable interest in the psychosocial aspects of communicative disorders. The literature is both extensive and wide-ranging, and much of it focuses on specific types of impairment (e.g., stuttering, language impairment). There are two general areas of study, however, that are of particular interest. The first is the frequent co-occurrence of speech and language impairment and socioemotional problems. A great deal of research has been directed toward exploring this relationship as well as toward determining what mechanisms might underlie this comorbidity. A second area of interest concerns the long-term outcomes of communicative problems across various areas of psychosocial development (e.g., peer relations, socioemotional status). Both of these lines of work are briefly discussed here.

Co-occurrence of Disorders. Numerous investigators have reported a high level of co-occurrence between communicative disorders and socioemotional problems. This high level of co-occurrence has been observed in various groups of children, including both those with a primary diagnosis of speech and language impairment and those with a primary diagnosis of psychiatric impairment or behavior disorder. Illustrative of these findings is the work of Baker and Cantwell (1987). These researchers performed psychiatric evaluations on 600 consecutive patients seen at a community speech, language, and hearing clinic. Children were divided into three subgroups of communication problems: speech (children with disorders of articulation, voice, and
fluency), language (children with problems in language expression, comprehension, and pragmatics), and a speech and language group (children with a mixture of problems). Of these children, approximately 50% were diagnosed as having a psychiatric disorder. These problems were categorized into two general groups of behavior disorder and emotional disorder.

Several researchers have speculated on the basis for this high level of co-occurrence between communication and socioemotional disorders. For example, Beitchman, Brownlie, and Wilson (1996) proposed several potential relationships, including the following: (1) impaired communicative skills lead to socioemotional impairment, (2) impaired communicative skills result in academic problems, which in turn lead to behavioral problems, (3) other variables (e.g., socioeconomic status) explain, in part or in whole, the relationship between communicative problems and socioemotional difficulties, and (4) an underlying factor (e.g., neurodevelopmental status) accounts for both types of problems.

Further research is needed to clarify the relationship between speech and language ability and socioemotional status. One approach to this problem has been to investigate various child factors that may contribute to developmental risk. For example, Tomblin et al. (2000) reported that reading disability is a key mediating factor predicting whether children with language impairment demonstrate behavioral difficulties.

Of particular interest is the relationship between social competence, communicative competence, and socioemotional functioning. It is clear that speech and language skills play a critical role in social interaction and that children who have difficulty communicating are likely to have difficulty interacting with others. The way in which various components of behavior interact, however, is not as straightforward as might initially be thought. For example, Fujiki et al. (1999) found that children with language impairment were more withdrawn and less sociable than their typical peers, consistent with much of the existing literature. More specific evaluation revealed that these differences were based on particular types of withdrawal (reticence, solitary active withdrawal). Further, severity of language impairment, at least as measured by a formal test of language, was not related to severity of withdrawal. Further clarification is needed to determine how these areas of development interact to produce social outcomes, and what factors may exacerbate or moderate socioemotional status.

Long-Term Consequences. A related line of work has focused on the long-term psychosocial and sociobehavioral consequences of speech and language impairment. In summarizing numerous studies looking at the outcomes of communication disorders, Aram and Hall (1989) stated that children with language impairment have frequently been found to have high rates of persistent social and behavioral problems. Children with speech impairment tend to have more favorable long-term outcomes.

The work of Beitchman and colleagues provides one example of a research program examining long-term psychosocial outcomes of individuals with communicative impairment. These researchers followed children with speech impairment and language impairment and their typical controls longitudinally over a 14-year period (Beitchman et al., 2001). At age 5, the children in the group with speech impairment and the group with language impairment had a higher rate of behavioral problems than the control group. At age 12, socioemotional status was closely linked to status at age 5. At age 14 years and at age 19 years, individuals in the group with language impairment had significantly higher rates of psychiatric involvement than the control group. Children in the group with speech impairment did not differ from the controls.

A few studies have examined the long-term psychosocial outcomes of individuals with speech and/or language impairment as they enter adulthood. For example, Records, Tomblin, and Freese (1992) examined quality of life in a group of 29 young adults (mean age, 21.6 years) with specific language impairment and 29 controls. The groups did not significantly differ on reported personal happiness or life satisfaction. Additionally, differences were not observed with respect to satisfaction in relation to specific aspects of life, such as employment or social relationships.

Howlin, Mawhood, and Rutter (2000) reported a bleaker picture. They reexamined two groups of young men, 23–24 years of age, who had first been evaluated at 7–8 years of age. One group was identified with autism and the other with language impairment. At follow-up, the group with language impairment showed fewer social and behavioral problems than the group with autism. The two groups had converged over the years, however, and differences between the two were not qualitative. The young men with language impairment showed a high incidence of social difficulties, including problems with social interaction, limited social contacts, and difficulty establishing friendships. Most still lived with their parents and had unstable employment histories in manual or unskilled jobs. Neither childhood language ability nor current language ability predicted social functioning in adulthood. Howlin et al. (2000) concluded that in language impairment, “as in autism, a broader deficit underlies both the language delay and the social impairments” (p. 573).

Given some of the data cited above, it would appear that children with speech difficulties achieve better psychosocial outcomes than children with language difficulties (see also Toppelberg and Shapiro, 2000). Although this may generally be the case, generalizations across individuals with different types of speech impairment must be made with caution. Some types of speech problems, such as stuttering, are likely to have important psychosocial implications, but also have relatively low incidence rates. Thus, in large group design studies where individuals are categorized together under the general heading of “speech,” the unique psychosocial difficulties associated with such disorders may be masked.
by the psychosocial profiles associated with more commonly occurring communication problems.

It should also be noted that speech impairments may vary from having no outward manifestations aside from those involved in talking to relatively severe physical or cognitive deficits. The impact of associated problems on the psychosocial development of children with differing types of communicative impairment is difficult to summarize briefly. Illustrative of the complexity even within a specific category of speech impairment are children with cleft lip and palate. These children may have articulation problems and hypernasality secondary to specific physical anomalies. These physical anomalies may be resolved, to various degrees, with surgery. Speech may also vary considerably. No specific personality type has been associated with children with cleft palate (Richman and Eliason, 1992). Individual studies, however, have found these children to exhibit higher than expected rates of both internalizing and externalizing behavior (Richman and Millard, 1997). It appears that factors such as family support, degree of disfigurement, and self-appraisal interact in complex ways to produce psychosocial outcomes in children with cleft lip and palate.

In summary, it is clear that individuals with communicative disorders often have difficulty with aspects of psychosocial behavior and that these problems can have long-term implications. There is also evidence that children with language impairment have more psychosocial difficulties than children with speech impairment. It must be remembered, however, that speech problems differ by type of impairment, severity, and other variables. Thus, generalizations must be made with caution. Given the accumulated evidence, there is good reason to believe that parents, educators, and clinicians working with children with speech and language impairment should give serious consideration to psychosocial status in planning a comprehensive intervention program.

See also Poverty: Effects on Language; Social Development and Language Impairment.

—Martin Fujiki and Bonnie Brinton

References


Further Readings


**Speech and Language Disorders in Children: Computer-Based Approaches**

Computers can be used effectively in the assessment of children’s speech and language. Biofeedback instrumentation allows the clinician to obtain relatively objective measures of certain aspects of speech production. For example, measures of jitter and shimmer can be recorded, along with perceptual judgments about a client’s pitch and intensity perturbations (Case, 1999). Acoustic analyses (Kent and Read, 1992) can be used to supplement the clinician’s perceptions of phonological contrasts (Masterson, Long, and Buder, 1998). For the evaluation of a client suspected of having a fluency disorder, recent software developments allow the clinician to gather measures of both the number and type of speech disfluencies and to document signs of effort, struggle, or disruption of airflow and phonation (Baker, 1999a). Hallowell (1999) discusses the use of instrumentation for detecting and measuring eye movements for the purpose of comprehension assessment. This exciting tool allows the clinician to evaluate comprehension in a client for whom traditional response modes, such as speaking or even pointing, are not possible.

Computers can also be used to administer or score a formal test (Cochran and Masterson, 1995; Hallowell and Katz, 1999; Long, 1999). Computer-based scoring systems allow the input of raw scores, which are then converted to profiles or derived scores of interest (Long, 1999). The value of such programs is inversely related to the ease of obtaining the derived scores by hand. If the translation of raw scores to derived scores is tedious and time-consuming, clinicians might find the software tools worth their investment in time and money.

Although few computerized tests are currently available, the potential for such instruments is quite high. Hallowell and Katz (1999) point out that computerized test administration could allow tighter standardization of administration conditions and procedures, tracking of response latency, and automated interfacing with alternative response mode systems. Of particular promise are the computerized tests that adapt to a specific client’s profile. That is, stimuli are presented in a manner that is contingent on the individual’s prior responses (Letz, Green, and Woodard, 1996). The type of task or specific items that are administered can be automatically determined by a client’s ongoing performance (e.g., Masterson and Bernhardt, 2001), which makes individualized assessment more feasible than ever. Incorporation of some principles from artificial intelligence also makes the future of computers in assessment exciting. For example, Masterson, Apel, and Wasowicz (2001) developed a tool for spelling assessment that employs complex algorithms for parsing spelling words into target orthographic structures and then aligning a student’s spelling with the appropriate correct forms. Based on the type of misspellings exhibited by each individual student, the system identifies related skills that need testing, such as phonological awareness or morphological knowledge. This system makes possible a comprehensive description of a student’s spelling abilities that would otherwise be prohibitive because of the time required to perform the analyses by hand and administer the individualized follow-ups.

Computerized language and phonological sample analysis (CL/PSA) has been in use since the 1980s (Evans and Miller, 1999; Long, 1999; Masterson and Oller, 1999; Long and Channell, 2001). These programs allow researchers and clinicians to perform complex, in-depth analyses that would likely be impossible without the technology. They provide instant analysis of a wide range of phonological and linguistic measures, and some provide tools that reduce and simplify the time-consuming process of transcribing samples (Long, 1999). Many of the CL/PSA programs also include comparison databases of language samples from both typical and clinical populations (Evans and Miller, 1999). Despite the power of CL/PSA programs, their use in clinical settings remains limited, for unclear reasons. It is possible that funding for software and hardware is insufficient; however, data from recent surveys (McRay and Fitch, 1996; ASHA, 1997) do not support this conjecture, since most
respondents do report owning and using computers for other purposes. Lack of use is more likely related to insufficient familiarity with many of the measures derived from language sample analysis and failure to recognize the benefits of these measures for treatment planning (Cochran and Masterson, 1995; Fitch and McRay, 1997). In an effort to address this problem, Long established the Computerized Profiling Website (http://www.computerizedprofiling.org) in 1999. Clinicians can visit the web site and obtain free versions of this CL/PSA software as well as instructional materials regarding its use and application.

Computer software for use in speech and language intervention has progressed significantly from the early versions, which were based primarily on a drill-and-practice format. Cochran and Nelson (1999) cite literature that confirms what many clinicians knew intuitively: software that allows the child to be in control and to independently explore based on personal interests is more beneficial than computer programs based on the drill-and-practice model. Improvements in multimedia capacities and an appreciation for maximally effective designs have resulted in a proliferation of software packages that can be effectively used in language intervention with young children. As with any tool, the focus must remain on the target linguistic structures rather than the toys or activities that are used to elicit or model productions. In addition to therapeutic benefits, computers offer reasonable compensatory strategies for older, school-age students with language-learning disabilities (Wood and Masterson, 1999; Masterson, Apel, and Wood, 2002). For example, word processors with text-to-speech capabilities allow students to check their own work by listening to as well as reading their text. Spell and grammar checkers can be helpful, as long as students have been sufficiently trained in the optimal use of these tools, including an appreciation of their limitations. Speech recognition systems continue to improve, and perhaps someday they will free writers with language disorders from the burden of text entry, which requires choices regarding spelling, and spelling can be so challenging for students with language disorders that it interferes with text construction. Currently, speech recognition technology remains limited in recognition accuracy for students with language disorders (Wetzel, 1996). Even when accuracy improves to an acceptable level, students will still need specific training in the optimal use of the technology. Optimal writing involves more than a simple, direct translation of spoken language to written form. Students who employ speech recognition software to construct written texts will need focused instruction regarding the differences between the styles of spoken and written language. Finally, the Internet provides not only a context for language intervention, but a potential source of motivation as well. The percentage of school-age children who use the Internet on a daily basis for social as well as academic purposes continues to increase, and it is likely that speech-language pathologists will capitalize on this trend.

Computers add a new twist to an old standard in phonological treatment. Instead of having to sort and carry numerous picture cards from one treatment session to the next, clinicians can choose one of several software packages that allow access and display of multimedia stimuli on the basis of phonological characteristics (Masterson and Rvachew, 1999). New technologies, such as the palatometer, provide clients with critical feedback for sound production when tactile or kinesthetic feedback has not been sufficient. Similarly, computer programs can be used to provide objective feedback regarding the frequency of stutterings, which might be considered less confrontational than feedback provided by the clinician (Bakker, 1999b). One particularly promising technology, the Speech Enhancer, incorporates real-time processing of an individual's speech production and selectively boosts energy only in those frequencies necessary for maximum intelligibility. Cariski and Rosenbek (1999) collected data from a single subject and found that intelligibility scores were higher when using the Speech Enhancer than when using a high-fidelity amplifier. The authors suggested that their results supported the notion that the device did indeed do more than simply amplify the speech output.

The decision to use computers in both assessment and treatment activities will continue to be based on the clinician's judgment as to the added value of the technology application. If a clinician can do an activity just as well without a computer, it is unlikely that she or he will go to the expense in terms of time and money to invest in the computer tool. On the other hand, for those tasks that cannot be done as well or even at all, clinicians will likely turn to the computer if they are convinced that the tasks themselves are worth it.

See also APHASIA TREATMENT: COMPUTER-AIDED REHABILITATION.

—Julie J. Masterson

References


**Further Readings**

Speech and Language Issues in Children from Asian-Pacific Backgrounds

Asian-Pacific Americans originate from Pacific Asia or are descendants of Asian-Pacific island immigrants. Numbering 10,477,000 in the United States, Asian-Pacific Americans are the fastest-growing segment of the U.S. population, representing 3.8% of the nation’s population and 10% of California’s population (Population Reference Bureau, 2001). By the year 2020, Asian-American children in U.S. schools will total about 4.4 million.

The recent Asian influx represents a diverse group from Southeast Asia, China, India, Pakistan, Malaysia, Indonesia, and other Pacific Rim areas. In general, Pacific Asia is divided into the following regions: East Asia (China, Taiwan, Japan, and Korea), Southeast Asia (Philippines, Vietnam, Cambodia, Laos, Malaysia, Singapore, Indonesia, Thailand), the Indian subcontinent, or South Asia (India, Pakistan, Bangladesh, Sri Lanka), and the Pacific islands (Polynesia, Micronesia, Melanesia, New Zealand, and Australia). Asian-Pacific populations speak many languages, and their English is influenced by various dialects and languages.

Asian-Pacific Americans are extremely diverse in all aspects of life, including attitudes toward disability and treatment, childrearing practices, languages, and culture. The Asian-Pacific island cultures, however, have interacted with and influenced each other for many generations, and therefore share many similarities. The following information is presented to provide an understanding of Asian-Pacific Americans in order to assist speech-language pathologists and audiologists in providing services to these culturally and linguistically diverse populations. Recommended assessment procedures and intervention strategies are provided.

Attitudes Toward Disability and Treatment Methods. What constitutes a disability depends on the values of the cultural group. In general, Eastern cultures may view a disabling condition as the result of wrongdoing of the individual’s ancestors, resulting in guilt and shame. Disabilities may be explained by a variety of spiritual or cultural beliefs, such as an imbalance in inner forces, bad karma (fate) or a curse. All over the world, people use different methods to treat illnesses and diseases, including consulting with priests, healers, herbalists, Qi-Gong specialists, clansmen, shamans, elders, and physicians. Among the Hmong, for example, surgical intervention is viewed as invasive and harmful.

Childrearing Practices. Childrearing practices and expectations of children vary widely from culture to culture (Westby, 1990; Van Kleek, 1994). There are differences in how parents respond to their children’s language, who interacts with children, and how parents and families encourage children to initiate and continue a verbal interaction. However, socioeconomic and individual differences must always be considered.

Languages. The hundreds of different languages and dialects that are spoken in East and Southeast Asia and the Pacific islands can be classified into five major families: (1) Malayo-Polynesian (Austronesian), including Chamorro, Ilocano, and Tagalog; (2) Sino-Tibetan, including Thai, Yao, Mandarin, and Cantonese; (3) Austro-Asiatic, including Khmer, Vietnamese, and Hmong; (4) Papuan, including New Guinean; and (5) Altaic, including Japanese and Korean (Ma, 1985). Additionally, there are 15 major languages in India from four language families, Indo-Aryan, Dravidian, Austro-Asiatic, and Tibeto-Burman (Shekar and Hegde, 1995).

Cultural Tendencies. Cultural tendencies of Asian-Pacific Americans may be quite different from those of individuals born and raised in a Western culture. Table 1 provides a sampling of these differences. However, caution should be taken not to overgeneralize this information in relation to a particular client or family.

**Table 1. Cultural Differences Between Asian-Pacific Americans and Western Groups**

<table>
<thead>
<tr>
<th>Eastern Tendencies</th>
<th>Western Tendencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>A person is not autonomous.</td>
<td>A person is autonomous.</td>
</tr>
<tr>
<td>A person is part of society.</td>
<td>A person is unique and individualistic.</td>
</tr>
<tr>
<td>A person needs to maintain relationships and have constraints.</td>
<td>A person makes rational choices.</td>
</tr>
<tr>
<td>A person is oriented toward harmony.</td>
<td>A person is active in decision making.</td>
</tr>
<tr>
<td>A person is a partner in the community where people are mutually responsible for behaviors and consequences.</td>
<td>A person is responsible for own actions and takes the consequences.</td>
</tr>
<tr>
<td>A person needs to be humble, improve, and master skills.</td>
<td>A person needs to feel good about self.</td>
</tr>
<tr>
<td>A person needs to endure hardships and persevere.</td>
<td>A person needs to toot own horn.</td>
</tr>
<tr>
<td>A person needs to self-reflect.</td>
<td></td>
</tr>
</tbody>
</table>

Recommended Assessment Procedures

The following assessment guidelines are often referred to as the RIOT (review, interview, observe, test) protocol (Cheng, 1995, 2002). They are adapted here for Asian-Pacific American populations:

1. **Review** all pertinent documents and background information. Many Asian countries do not have medical records or cumulative school records. Oral reports are sometimes unreliable. A cultural informant or an interpreter is generally needed to obtain this information because of the lack of English language proficiency of the parents or guardians. Pregnancy and delivery records...
might not have been kept, especially if the birth was a home birth or in a refugee camp.

2. Interview teachers, peers, family members, and other informants and work with them to collect data regarding the client and the home environment. The family can provide valuable information about the communicative competence of the client at home and in the community, as well as historical and comparative data on the client’s language development. The clinician needs information regarding whether or not the client is proficient in the home language. The family’s home language, its proficiency in different languages, the patterns of language usage, and the ways the family spends time together are some areas for investigation. Interview questions are available from multiple sources (Cheng, 1990, 1991, 2002; Langdon and Saenz, 1996). Questions should focus on obtaining information on how the client functions in his or her natural environment in relation to age peers who have had the same or similar exposure to their home language or to English.

3. Observe the client over time in multiple contexts with multiple communication partners. Observe interactions at school, both in and outside the classroom, and at home. This cognitive-ecological-functional model takes into account the fact that clients often behave differently in different settings (Cheng, 1991). Direct observation of social behavior with multiple participants allows the evaluator to observe the ways members of different cultures view their environment and organize their behavior within it.

4. Test the client using informal and dynamic assessment procedures in both the school language and the home language. Use the portfolio approach by keeping records of the client’s performance over time. Interact with the client, being sensitive to his or her needs to create meaning based on what is perceived as important, the client’s frame of reference, and experiences.

What clinicians learn from the assessment should be integrated into their intervention strategies. Intervention should be constructed based on what is most productive for promoting communication and should incorporate the client’s personal and cultural experiences. Salient and relevant features of the client’s culture should be highlighted to enhance and empower the client.

There are many challenges professionals face in working with the Asian-Pacific American populations. The discourse styles in effect in American homes and schools may differ from those that are practiced in Asian classrooms. Clinicians need to be doubly careful and not interpret these differences as deficient, disordered, aberrant, and undesirable. Behaviors that can be easily misunderstood include the following (Cheng, 1999):

- Delay or hesitation in response
- Frequent topic shifts and poor topic maintenance
- Confused facial expressions, such as a frown signaling concentration rather than displeasure
- Short responses
- Use of a soft-spoken voice
- Taking few risks
- Lack of participation and lack of volunteering information
- Different nonverbal messages
- Embarrassment over praise
- Different greeting rituals, which may appear impolite, such as looking down when the teacher approaches
- Use of Asian-language-influenced English, such as the deletion of plural and past tense

These are just a few examples of the observed behaviors that may be misinterpreted. Asian-Pacific American children may be fluent in English but use the discourse styles of their home culture, such as speaking softly to persons in authority, looking down or away, and avoiding close physical contact. Surface analysis of linguistic and pragmatic functions is not sufficient to determine the communicative competence of children and might even misguide the decision-making process.

Sociocultural and psychological difficulties arise in the conflict of culture, language, and ideology between Asian students, their parents, and the American educational system. These difficulties can include the background of traditions, religions, and histories of the Asian-Pacific population, problems of acculturation, the understanding of social rules, contrasting influences from home and the classroom, confusion regarding one’s sense of identity relating to culture, society, and family, the definition of disability, and the implications of receiving special education services.

Intervention activities and materials can be selected based on the client’s family and cultural background, using activities that are culturally and socially relevant. In addition to traditional intervention techniques of modeling and expansion, speech-language pathologists can include activities such as those discussed by Cheng (1989).

Alternative strategies should be offered when clients or caregivers are reluctant to accept the treatment program recommended by the speech-language pathologist or audiologist. Inviting them to special classes or speech and language sessions is a useful way to provide the needed information. Seeking assistance from community leaders and social service providers may also be necessary to convince the clients of the importance of therapy or recommended programs. The clients or caregivers may also be asked to talk with other Asian-Pacific Americans who have experiences with treatment programs. Other individuals can be effective in sharing their personal stories about their experiences with therapy. The clinician should be patient with the clients, letting them think through a problem and waiting for them to make the decision to participate in the treatment program.

Following are some suggestions to create an optimal language learning environment and to reduce difficult communication (Cheng, 1996):

- Make no assumptions about what students know or do not know.
- Anticipate their needs and greatest challenges.
- Expect frustration and possible misunderstanding.

...


- Encourage students to join social activities such as student government, clubs, and organizations to increase their exposure to different types of discourse, as language is a social tool and should be used for fulfilling multiple social needs and requirements.
- Facilitate the transition into mainstream culture through such activities as role-playing (preparing scripts for commonly occurring activities, using culturally unique experiences as topics for discussions) and conducting social/pragmatic activities (Cheng, 1989), such as a birthday party and a Thanksgiving celebration.
- Nurture bicultural/multicultural identity. Introduce multicultural elements not only in phonology, morphology, and syntax, but also in pragmatics, semantics, and ritualized patterns.

Providing speech-language and hearing services to Asian-Pacific Americans is challenging. Preassessment information on the language, culture, and personal history of the individual lays a solid foundation to further explore the client’s strengths and weaknesses. Assessment procedures need to be guided by the general principles of being fair to the culture and nonbiased. The results of assessment should take into consideration the cultural and pragmatic variables of the individual. Intervention can be extremely rewarding when culturally relevant and appropriate approaches are used. The goals of intervention must include the enhancement of appropriate language and communication behaviors, home language, and literacy. Clinicians need to be creative and sensitive in their intervention to provide comfortable, productive, and enriching services for all clients. Some publishers that have developed materials for use with Asian-Pacific American populations include Academic Communication Associates (Oceanside, CA), Communication Skill Builders (Tucson, AZ), Thinking Publications (Eau Claire, WI), Newbury House (Rowley, MA).

—Li-Rong Lilly Cheng

References


On-line Resources

http://www.krysatal.com/langfams.html
http://www.travlang.com/languages
http://www.zompist.com/lang8.html

Speech Assessment, Instrumental

The instrumental analysis of speech can be approached through the three main stages of the speech chain: articulatory, acoustic, and auditory phonetics. This entry reviews the instrumentation used to assess the articulatory and acoustic phases. Auditory phonetic techniques are covered elsewhere in this volume.

Although speech planning in the brain (neurophonetics) lies outside the traditional tripartite speech chain, the neurological aspects of both speech production and perception can be studied through the use of brain imaging techniques. Speech articulation proper is deemed here to begin with the movement of muscles required to produce aerodynamic changes resulting in the flow of an airstream (see Laver, 1994; Ball and Rahilly, 1999). Of course, muscle movements also occur throughout articulation. This area has been investigated using electromyography (EMG). In EMG, electrodes of different types (surface, needle, and hooked wire) are used to gather data on electrical activity within target muscles, and these data are matched with a simultaneously recorded speech signal (Stone, 1996; Gentil and Moore, 1997). In this way the timing of muscle activity in relation to different aspects of speech can be investigated. This technique has been used to examine both normal and disordered speech. Areas studied include...
the respiratory and laryngeal muscles, muscle groups in the lips, tongue, and soft palate, and various disorders, including disorders of voice and fluency, and certain acquired neurological problems. Figure 1 shows EMG traces from a patient with Friedreich’s ataxia.

Aerodynamic activity in speech is studied through aerometry. A variety of devices have been used to measure speech aerodynamics (Zajac and Yates, 1997). Many systems have employed an airtight mask that is placed over the subject’s face and attached to a pneumotachograph. The mask contains sensors to measure pressure changes and airflow at the nose and mouth, and generally also a microphone to record the speech signal, against which the airflow can be plotted. If the focus of attention is lung volume changes, then a plethysmograph may be employed. This is an airtight box that houses the subject, and any changes to the air pressure within the box (caused by changes in the subject’s lung volume) are recorded. A simpler plethysmograph (the respitrace; see Stone, 1996) consists of a wire band placed around the subject’s chest that measures changes in cross-sectional area during inhalation and exhalation.

In normal pulmonic egressive speech, airflow from the lungs passes through the larynx, where a variety of phonation types may be implemented. The study of laryngeal activity (more particularly, of vocal fold activity) can be direct or indirect. In direct study, a rigid or flexible endoscope connected to a camera is used to view the movements of the folds. This technology is often coupled to a stroboscopic light source, as stroboscopic endoscopy allows the viewer to see individual movements of the folds (see Abberton and Fourcin, 1997). Endoscopy, however, is invasive, and use of a rigid endoscope precludes normal speech. Indirect investigation of vocal fold activity is undertaken with electroglottography (EGG), also termed electrolaryngography (see Stone, 1996; Abberton and Fourcin, 1997). This technique allows vocal fold movement to be extrapolated from measuring the varying electrical resistance across the larynx. Both approaches have been used in the investigation of normal and disordered voice.

Velic action and associated differences in oral and nasal airflow (and hence in nasal resonance) can also be measured directly or indirectly. The velotrace is an instrument designed to indicate directly the height of the velum (see Bell-Berti et al., 1993), while nasometric devices of varying sophistication measure oral versus nasal airflow (see Zajac and Yates, 1997). The velotrace is invasive, as part of the device must be inserted into the nasal cavity to sit on the roof of the velum. Nasometers measure indirectly using, for example, two external microphones to measure airflow differences. Figure 2 shows a trace from the Kay Elemetrics nasometer of hypernasal speech.

The next step in the speech production chain is the articulation of sounds. Most important here is the placement of the individual articulators, and electro-palatography (EPG) has proved to be a vital development in this area of study. Hardcastle and Gibbon

Figure 1. Averaged integrated EMG signals for the mentalis muscle (MENT), orbicularis oris inferior (OOI), orbicularis oris superior (OOS), anterior belly of the digastric (ABD), and the depressor labii inferior (DLI) for a patient with Friedreich’s ataxia uttering /epapap/. (Courtesy of Michèle Gentil.)

Figure 2. Trace adapted from a Kay Elemetrics nasometer showing normal and hypernasal versions of “eighteen, nineteen, twenty.”
(1997) describe this technique. A thin acrylic artificial palate is made to fit the subject. This palate has a large number of electrodes embedded in it (from 62 to 96, depending on the system employed) to cover important areas for speech (e.g., the alveolar region). When the tongue touches these electrodes, they fire, and the resultant tongue-palate contact patterns can be shown on a computer screen. The electrodes are normally sampled 100 times per second, and the patterns are displayed in real time. This allows the technique to be used both for research and for feedback in therapy. EPG has been used to study normal speech and a wide range of disordered speech patterns. Figure 3 shows tongue-palate contact patterns in a stylized way for two different EPG systems.

Other ways of examining articulation (and indeed a whole range of speech-related activity) can be subsumed under the overall heading of speech imaging (see Stone, 1996; Ball and Gröne, 1997). The oldest of these techniques is x-radiography. A variety of different x-ray techniques have been used in speech research, among them videofluorography, which uses low doses of radiation to give clear pictures of the vocal tract, and x-ray microbeam imaging, in which the movements of pellets attached to relevant points of the tongue and palate are tracked. Because of the dangers of radiation, alternative imaging techniques have been sought. Among these is ultrasound, which uses the time taken for sound waves to bounce off a structure and return to a receiver to map structures in the vocal tract. Because ultrasound waves do not travel through the air, mapping of the tongue (from below) is possible, but mapping of tongue–palate distances is not, as the palate cannot be mapped through the air space of the oral cavity. Electromagnetic articulography (EMA) is another tracking technique. In this technique the subject is placed within alternating magnetic fields generated by transmitter coils in a helmet assembly. Small receiver coils are placed at articulatorily important sites (e.g., tongue tip, tongue body). The movements of the receiver coils through the alternating magnetic fields are measured and recorded by computer. As with x-ray microbeam imaging, the tracked points can be used to infer the shape and movements of articulators within the vocal tract.

The final imaging technique to be considered is magnetic resonance imaging (MRI). The imager surrounds a subject with electromagnets, creating an electromagnetic field. This field causes hydrogen protons (abundant in human tissue) to align but also to precess, or wobble. If a brief radio pulse is introduced at the same frequency as the precessing, the protons are moved out of alignment and then back again. As they realign, they emit weak radio signals, which can be used to construct an image of the tissue involved. MRI can provide good images of the vocal tract but currently not at sufficient frequency to allow analysis of continuous speech. All of these imaging techniques have been used to study aspects of both normal and disordered speech. Figure 4 shows ultrasound diagrams for two vowels and two consonants.

Acoustic analyses via sound spectrography are now easily undertaken with a range of software programs on personal computers as well as on dedicated hardware-software configurations such as the Kay Elemetrics Sonagraph. Reliable analysis depends on good recordings (see Tatham and Morton, 1997). Spectrographic analysis packages currently allow users to analyze temporal, frequency, amplitude, and intensity aspects of a speech signal (see Baken and Daniloff, 1991; Farmer, 1997). For example, a waveform displays amplitude versus time; a wideband spectrogram displays frequency versus time using a wideband pass filter (around 200–300 Hz), giving good time resolution but poor frequency resolution; a narrow-band spectrogram shows frequency versus time using a narrowband pass filter (around 29 Hz), which provides good frequency resolution but poor time resolution; and spectral envelopes show frequency versus intensity at a point in time (produced either by fast Fourier transform or linear predictive coding). Speech analysis research has generally concentrated on wideband spectrograms and spectral envelopes. These
Figure 4. Ultrasound images of two vowels and two consonants. (Courtesy of Maureen Stone.)

Figure 5. Wideband spectrogram of a disfluent speaker producing “(provin)c(ial t(owns).” (Courtesy of Joan Rahilly.)
both show formant frequencies (bands of high intensity at certain frequency levels), which are useful in the identification of and discrimination between vowels and other sonorants. Fricatives are distinguishable from the boundaries of the broad areas of frequency seen clearly on a spectrogram, while plosives can be noted from the lack of acoustic activity during the closure stage and the coarticulatory effects on the formants of neighboring sounds. Segment duration can easily be measured from spectrograms in modern analysis packages. Various pitch extraction algorithms are provided for the investigation of intonation. Farmer (1997) provides an extensive review of acoustic analysis work in a range of sounds. Segment duration can easily be measured from spectrograms in modern analysis packages. Various pitch extraction algorithms are provided for the investigation of intonation. Farmer (1997) provides an extensive review of acoustic analysis work in a range of sounds.

—Martin J. Ball

References


Further Readings

Speech Assessment in Children: Descriptive Linguistic Methods

Descriptive linguistic methods have long been used in the analysis of fully developed primary languages. These same methods are also well-suited to the study of language development, particularly the analysis of children’s speech sound systems. Descriptive methods are a preferred analytic tool because they are designed to gather evidence that reveals the hallmark and defining characteristics of a sound system, independent of theoretical orientation, age, or population of study. The defining properties of descriptive linguistic analyses of children’s sound systems are discussed in this article.

The Phonemic Inventory. In development, phonemic gaps in the inventory property Y is unmarked and predictably easier to learn. In children, then, phonemic gaps in the inventory correspond to more marked (difficult) structures of language. In linguistic terminology, these gaps would be characterized as a type of phonotactic constraint (Dinnsen, 1984).

The Distribution of Sounds. Distribution refers to where sounds (phones or phonemes) occur in words and is determined by examining context. For children, sounds may be used in all word positions, initial, intervocalic, and final, or they may be limited to certain contexts. In development, obstruent stops commonly occur word-initially but not postvocally; whereas fricatives and liquids commonly occur postvocally but not word-initially (Smith, 1973). As with the phonemic inventory, restrictions on the distribution of sounds correspond to markedness, with children having a tendency toward unmarked as opposed to marked structure.

Rule-Governed Alternations. Asymmetries in the distribution of sounds may be further indicative of systematic rule-governed alternations in sound production (Kenstowicz, 1994). Rule-governed alternations occur when morphologically related words are produced in different ways, for example, “electric” but “electricity.” Alternations are typically sampled by adding either a prefix or suffix to a base word in order to change the context in which a sound occurs. There are two general types of rule-governed change: allophonic variation and neutralization. Allophonic variation occurs when a single phoneme has multiple corresponding phonetic outputs that vary by context. An example is /t/ produced as aspirated in word-initial position “tap,” as flap in intervocalic position “bitter,” and as unreleased in word-final position “it.” In each case, the target sound is /t/, but the phonetic characteristics of the output differ predictably by word position. Thus, there is a one-to-many mapping between phoneme and phones in allophonic variation. Neutralization occurs when two or more phonemes are merged into one phonetic output in a well-defined context. An example is /t/ and /d/ both produced as flap in intervocalic position “writer” and “rider.” In neutralization, the contrast between phonemes is no longer apparent at the phonetic (surface) level. Consequently, there is a many-to-one mapping between phonemes and phone. In children, the emergence of target-appropriate morphophonemics occurs later in language development. For children with speech sound disorders, nontarget allophonic variation and neutralization have been observed and parallel the rules of fully developed languages of the world (Camarata and Gandour, 1984).

Together, these four properties define the most basic elements of a sound system at a segmental level of structure. In addition to examining these properties, descriptive linguistic methods may evaluate prosodic levels of structure by examining units larger than the sound, such as permissible syllable types and combinations and the overlay of primary and secondary stress on these in the formation of words and phrases (Liéo and Prinz, 1996; Kehoe and Stoel-Gammon, 1997). As with segmental structure, children typically use unmarked pro-
sodic structure, with preferences for open syllables and trochaic (strong-weak) stress assignment.

For children with speech sound disorders, there are other methods of analysis that may be relevant to a comprehensive characterization of the sound system (Fey, 1992; see SPEECH SOUND DISORDERS IN CHILDREN: DESCRIPTION AND CLASSIFICATION). Relational analyses establish a one-to-one correspondence between a child’s errored outputs and intended target sounds. These analyses are intended to capture the patterns of a child’s errors, and to descriptively label these patterns as phonological processes. Four main categories of phonological processes characterize children’s commonly occurring developmental errors (Ingram, 1989). These categories are substitution processes, involving different manners or places of production than the target; syllable shape processes, involving different canonical (consonant-vowel) shapes than the target; assimilatory processes, involving sounds produced more alike in a word than in the target; and other processes, such as reversals in the sequencing of sounds or articulatory differences in sound production such as lisping. Children with speech sound errors are likely to use other unusual phonological processes and to persist in their use of these processes for longer durations than are typical (Leonard, 1992).

Supplementary clinical methods have also been designed to evaluate perceptual or metalinguistic skills, as these skills may affect a child’s knowledge of the ambient sound system. The Speech Production-Perception Task is one clinical technique that establishes a child’s ability to perceptually differentiate target sounds from their corresponding substitutes (Locke, 1980). Other metalinguistic procedures employ categorization tasks that evaluate a child’s judgment of the similarity of target sounds and their substitutes (Klein, Lederer, and Cortese, 1991). Although these methods may have clinical utility in isolating the source of breakdown and in designing appropriate intervention for a child’s speech disorder, they are considered external (not primary) evidence in conventional linguistic analyses of sound systems (Anderson, 1981), because these skills lie outside the domain of phonology in particular and language in general.

Finally, one of the most central aspects of a descriptive linguistic analysis of a sound system is the interpretation or theoretical account of the data. A number of theories have been advanced to account for the fundamental properties of sound systems. Each relies on a unique set of assumptions about the structure, function, and organization of sounds in a speaker’s mental lexicon. Among the most recognized frameworks are linear phonology, including standard generative and natural frameworks; nonlinear phonology, including autosegmental, metrical, underspecification, and feature geometry frameworks (Goldsmith, 1995); and, most recently, optimality theory (Prince and Smolensky, 1997). Any formal theory of language must account for the facts of acquisition, including those pertinent to children with speech sound disorders. Acquisition data present unique challenges for linguistic theory because of the inherent variability within and across children’s sound systems within and across points in time (Chomsky, 1999). These challenges have been handled in different ways by different linguistic theories, but at the core, they have served to outline a well-defined set of research issues about children’s speech sound development that as of yet remain unresolved. Central questions bear on the nature of children’s mental (internal) representation of sound, the relationship between perception and production in speech sound development, and the contribution of innateness and maturation to language acquisition.

—Judith A. Gierut

References


**Further Readings**


**Speech Development in Infants and Young Children with a Tracheostomy**

A tracheostomy is a permanent opening of the trachea to outside air. It most often requires a surgical procedure for closure. The primary reason for performing a surgical tracheostomy is for long-term airway management in cases of chronic upper airway obstruction or central or obstructive sleep apnea, or to provide long-term mechanical ventilatory support. The use of assisted ventilation for more than 1 month in the first year of life has been considered to constitute a chronic tracheostomy (Bleile, 1993). Most of the estimated 900–2,000 infants and children per year who need a tracheostomy, a ventilator, or both for a month or more are, in fact, less than a year old (Singer et al., 1989). Although the mortality associated with a chronic tracheostomy in young children is twice that in adults, the procedure is invaluable for acute and long-term airway management (Fry et al., 1985).

When a tracheostomy or mechanical ventilation is used over a long period, the impact on communication and feeding behavior can be significant. Oral communication in children occurs in tandem with growth and maturation of the structures of the speech apparatus. Neuromuscular or biomechanical difficulties resulting from altered patterns of growth or structural problems can negatively affect the development of oral communication. In particular, the respiratory, laryngeal, and articulatory subsystems of the speech apparatus are at risk for pathological changes affecting the development of speech.

The respiratory subsystem of the speech apparatus comprises the lower airways, rib cage, diaphragm, and abdominal structures. The lower airways consist of the trachea, the right and left mainstem bronchi, and the lungs. The tracheobronchial tree is much smaller in children than in adults, and differs in shape and bio-
mechanics as well. The trachea in young children has been described as the size of a soda straw, and is highly malleable (Fry et al., 1985). The infant tracheal diameter is approximately 0.5 cm, whereas adult tracheal diameters are 1.5–2.5 cm. C-shaped cartilage rings joined by connective tissue help to keep the trachea from collapsing against the flow of air during breathing. Because the membranes of the infant trachea are soft and fragile, there is a risk of tracheal compromise secondary to a tracheostomy. Complications may include reactive granulation at the site of cannula, edema and scarring, chronic irritation of the tracheal lumen, and tracheal collapse as a result of increased negative pressure pulling air through a compromised structure (Fry et al., 1985).

During the first several years of life, significant changes occur in the structure and mechanics of the respiratory system. The airways increase in radius and length and the lungs increase in size and weight. The thoracic cavity enlarges and changes in shape, and overall chest wall compliance decreases with upright posture. Airway resistance decreases, and pleural pressure becomes more subatmospheric (Beckerman, Brouillette, and Hunt, 1992). Tidal volume, inspiratory capacity, vital capacity, and minute ventilation increase with age.

Besides providing phonation, the larynx, along with the epiglottis and soft palate, protects the lower airway. The infant larynx is located high in the neck, close to the base of the tongue. The thyroid cartilage is located directly below the hyoid bone, whereas the cricoid cartilage is the lowest part of the laryngeal structure. Because of its location and size, the infant larynx, like the trachea, is susceptible to trauma during airway management procedures. The laryngeal structures become less susceptible to injury as they change shape and descend during the first year of life.

The articulatory subsystem, composed of the pharynx, mouth, and nose, also undergoes significant changes in growth and function during infancy and early childhood. The pharynx plays a critical role in both respiration and swallowing. The infant pharynx lacks a rigid framework and can collapse if external suction is applied within the airway. If the airway-maintaining muscles are weak or paralyzed, normal negative pressures associated with inspiratory efforts also can cause airway collapse at the level of the pharynx (Thach, 1992). Movement of the pharyngeal walls, elevation of the soft palate, and elevation of posterior portion of the tongue are important maneuvers for achieving velopharyngeal closure. The infant tongue is proportionately larger in relation to mouth size than the adult’s; thus, tongue retraction can cause upper airway blockage and respiratory distress. Various craniofacial abnormalities may result in structural or neurological situations that require airway management interventions, including tracheostomy.

The decision to use long-term airway maintenance in the form of a tracheostomy requires consideration of many factors. Even the type of incision can make a difference in overall outcome for the infant or young child (Fry et al., 1985). Other information is needed to select the appropriate tracheostomy tube. Driver (2000) has compiled a list of the critical factors in tracheostomy tube selection. These factors include the child’s respiratory requirements, age and weight, tracheal diameter, distance from the tracheal opening to the carina, and anatomical features of the neck for selection of a neck plate or flange. In addition, decisions must be made regarding whether or not there should be an inner cannula, the flexibility of the cannula, whether or not there should be a cuff (an air-inflatable outer bladder used to create a seal against the outer wall of the tracheal tube and trachea), and what external adapters might be used (Driver, 2000).

Tracheostomy tubes are selected primarily on the basis of the ventilatory needs of the infant or young child. Tracheostomy tubes will be larger in diameter and may have a cuff in the event the child needs high ventilator pressures with frequent suctioning. However, because of a young child’s susceptibility to trauma of the speech apparatus, it is optimal to have a smaller-diameter, flexible tube without a cuff. When a smaller tube is selected, air leakage around the tube and through the upper airway will be available to the infant for voicing.

Many pediatric upper airway management problems can be successfully addressed with a tracheostomy alone. However, chronic respiratory failure will require some form of mechanical ventilation. The type of mechanical ventilation support required will depend on the type of disorder, the degree of respiratory dependence, and whether the child will ultimately be weaned from ventilatory support. There are two types of mechanical ventilation systems commonly used with children. Negative pressure ventilation is noninvasive and uses negative (below atmospheric) pressure by exerting suction on the outside of the chest and abdomen. As a result, intrathoracic pressure is reduced and induces airflow into the lungs. Expiration is accomplished by passive recoil of the lungs. Negative pressure ventilators work well with children who have relatively normal airways and compliant chest walls. Negative pressure ventilators are associated with fewer complications than positive pressure ventilators and do not require a tracheostomy (Splaingard et al., 1983). However, they are cumbersome and are not adequate for children with severe respiratory disease or rigid chest walls (Driver, 2000). Positive pressure ventilation is invasive and applies positive (above atmospheric) pressure to force air into the lungs via a ventilator connected to a tracheostomy tube (for long-term use). Expiration is accomplished by passive recoil of the lungs. The primary advantages are the flexibility to individualize respiratory support and to deliver various concentrations of oxygen (Metz, 1993). There are two major types of positive pressure ventilators, volume ventilators and pressure ventilators. In addition, ventilators are set in a mode (e.g., assist-control, synchronized intermittent mandatory ventilation) to deliver a certain number of breaths per minute, based on the tidal volume and minute ventilation.

Infants and young children with tracheostomies can become oral communicators if oral motor control is
sufficient, the velopharynx is competent, the upper airway is in reasonably good condition (i.e., there is no significant vocal fold paresis or paralysis and no significant airway obstruction), the ability to deliver airflow and pressure to the vocal folds and supporting laryngeal structures is sufficient, and chest wall muscular support for speech breathing is sufficient (not a prerequisite for ventilator-supported speech). If oral communication is possible, then the type of tracheostomy tube and the various valve configurations must be selected on the basis of both effective airway management and oral communication criteria. Driver (2000) suggests that the best results for oral communication are achieved if the smallest, simplest tracheostomy tube is selected. The size and nature (fenestrated vs. non-fenestrated) of the tracheostomy tube also will affect the effort required to move gas across the airway (Hussey and Bishop, 1996). Respiratory effort to breathe will impact on the additional effort required to vocalize and speak. The most efficient tracheostomy tube is one that has flow characteristics similar to those of the upper respiratory system for the maintenance of respiratory homeostasis, but with some trade-off for air leakage necessary for vocalization (Mullins et al., 1993). When tracheostomy tubes have a cuff, extreme caution should be taken to ensure that cuff deflation is accomplished prior to any attempts to support oral communication. Tracheostomy tubes with cuffs are not recommended for infants and very young children because of increased risk of tracheal wall trauma.

When a sufficient air leak around the tracheostomy tube exists, then a unidirectional speaking valve can be attached to the hub of the tube. When the child inspires, air enters through a diaphragm that closes on expiration, thus forcing air to exit through the upper airway. The same effect can be accomplished by manually occluding the hub of the tracheostomy tube. Speaking valves can be used with ventilator-assisted breathing as well. Several factors should be considered when selecting a pediatric speaking valve. These factors include the type of diaphragm construction (bias open or bias closed), the amount of resistance inherent in the valve type, and the amount of air loss during vocalization and speech production. First, speaking valves can be either bias open or bias closed at atmospheric pressure. A biased closed valve remains closed until negative air pressure is applied during inspiration. In this case, the valve will open during inspiration and close during expiration. A bias open valve remains open and only closes during the expiratory phase of the breath cycle. A bias closed valve may require greater effort to achieve airflow (Zajac, Frontaro-Clerici, and Roop, 1999). Differences in resistance have been found among valve types, especially during low flows (.450 liters/sec), however, all valves recently tested have resistances in the range of nasal resistance reported for normal adults. Whereas speaking valves have similar resistances, bias-open valves consistently show air loss during the rise in pressure associated with the /p/ consonant (Zajac, Frontaro-Clerici, and Roop, 1999).

Introducing a speaking valve to a young child can be challenging. The valve changes the sensation of breathing probably due to increases in resistance on both inspiration and expiration. Extra effort from expiratory muscles during phonation also must be generated to force air around the tracheostomy tube to the vocal folds. Finally, young children may not be familiar with coughing up secretions through the oral cavities and show distress until this skill is acquired (McGowan et al., 1993). Initially, the young child may be able to tolerate the speaking valve for only 5 minutes at a time. With encouragement and appropriate reinforcement, the child will likely tolerate the speaking valve for increasing amounts of time. When a speaking valve is placed in line with mechanical ventilation, various volume or pressure adjustments may be made to maximize the timing of phonation and the natural characteristics of the breathing and speaking cycles. Only a few general guidelines on mechanical ventilation and speech in infants and young children have been published (e.g., Lohmeier and Boliek, 1999).

A chronic tracheostomy interferes with the development of oral motor skills and experimentation with sound production by limiting movements of the jaw, tongue, and lips. In addition, long-term intubation may result in a significantly high-vaulted palate and vocal cord injury (Driver, 2000). Adequate breath support for speech also may be affected because of neuromuscular weakness, hypotonia, hypertonia, or paralysis. Consequently, infants and young children may have one or several issues affecting the speech mechanism. Infants and young children who need mechanical ventilation may not vocalize until near the end of the first year of life and may not be able to appropriately time their vocalizations to ventilator cycle until well after 12 months of age (Lohmeier and Boliek, 1999). The speech characteristics of infants and children with tracheostomies, using speaking valves or manual occlusion, include a smaller lung volume initiations, terminations, and excursions, fewer syllables per breath group, variable chest wall configurations during vocalization including rib cage or abdomen paradoxical movements, breathy or pressed voice quality that reflects available airflow and tracheal pressures, intermittent voice stops, hypernasality, and poor intelligibility (Lohmeier and Boliek, 1999). In addition, experimentation with vocal play, feeding, and oral-motor exploration may be limited during a sensitive period for speech and language acquisition. Therefore, all efforts should be made to support phonation and other communicative opportunities.

Only a handful of group and single case studies have assessed the developmental outcomes of speech and language following the long-term use of a tracheostomy or mechanical ventilation. These studies suffer from problems such as sample heterogeneity and the prelinguistic or linguistic status of the child at the time the intervention is performed, but they do suggest some general trends and outcomes. Fairly obviously, these children are at risk for delay in speech and language development (Simon and Handler, 1981; Simon, Fowler, and Handler, 1983; Kaslon and Stein, 1985; Simon and Mc-
Gowan, 1989; Bleile and Miller, 1994). Major gains in the development of speech and language can sometimes be made during and after decannulation with total communication intervention approaches. However, the data suggest that residual effects of long-term tracheostomy can be measured long after decannulation. Most reported delays seem to be articulatory in nature; voice and respiratory dysfunction are rarely reported in children after decannulation (Singer, Wood, and Lambert, 1985; Singer et al., 1989; Hill and Singer, 1990; Kamen and Watson, 1991; Kertoy et al., 1999). Taken together, these studies indicate possible residual effects of long-term tracheostomy on the speech mechanism. These effects appear unrelated to the time of intervention (i.e., prelinguistic or linguistic) but may be related to length of cannulation and the general constellation of medical conditions associated with long-term tracheostomy use.

—Carol A. Boliek

References


Further Readings


Speech Disfluency and Stuttering in Children

Childhood stuttering (also called developmental stuttering) is a communication disorder that is generally characterized by interruptions, or speech disfluencies, in the smooth forward flow of speech. Speech disfluencies can take many forms, and not all are considered to be atypical. Disfluencies such as interjections (“um”, “er”), phrase repetitions (“I want—I want that”), and revisions (“I want—I need that”), which are relatively common in the speech of normally developing children, represent normal aspects of the speaking process. These disfluencies arise when a speaker experiences an error in language formulation or speech production or needs more time to prepare a message. Other types of disfluencies, which occur relatively infrequently in the speech of normally developing children, may be indicative of a developing stuttering disorder. These disfluencies, often called “atypical” disfluencies, “stuttered” disfluencies, or “stutter-like” disfluencies, include whole-word repetitions (“I-I-I want that”) and, particularly, fragmentation within a word unit, such as part-word repetitions (“li-li-like this”), sound prolongations (“lllllike this”), and blocks (“I—iike this”) (Ambrose and Yairi, 1999).

Stuttered speech disfluencies can also be accompanied by affective, behavioral, and cognitive reactions to the difficulties with speech production. These reactions are distinct from the speech disfluencies themselves but are part of the overall stuttering disorder (Yaruss, 1998). Examples of behavioral reactions, which rapidly become incorporated into the child’s stuttering pattern, include physical tension and struggle in the speech mechanism as children attempt to control their speech. Affective and cognitive reactions include feelings of anxiety, embarrassment, and frustration. As stuttering continues, children may develop shame, low self-esteem, and avoidance of words, sounds, or speaking situations. These negative reactions can lead to increased stuttering severity and greatly exacerbate the child’s communication problems.

Etiology. Numerous theories about the etiology of stuttering have been proposed (Bloodstein, 1993). Historically, these theories tended to focus on single causes acting in isolation. Examples include psychological explanations based on supposed neuroses, physiological explanations involving muscle spasms, neurological explanations focusing on ticlike behaviors, and environmental explanations suggesting that normal disfluency was misidentified as stuttering. None of these theories has proved satisfactory, though, for the phenomenology of stuttering is complex and highly individualized. As a result, current theories focus on multiple etiological factors that interact in complex ways for different children who stutter (e.g., Smith and Kelly, 1997). These interactions involve not only genetic and environmental factors, but also various aspects of the child’s overall development.

Pedigree and twin studies have shown a genetic component to childhood stuttering—a family history of stuttering can be identified for approximately 60%–70% of children who stutter (Ambrose, Yairi, and Cox, 1993). The precise nature of that genetic inheritance is not fully understood, however, and studies are currently under way to evaluate different models of genetic transmission. It is also likely that environmental factors, such as the model the child hears when learning to speak and the
demands placed on the child to speak quickly or precisely, may play a role in determining whether stuttering will be expressed in a particular child (e.g., Starkweather and Givens-Ackerman, 1997).

There are several aspects of children's overall development that affect children's speech fluency and the development of childhood stuttering. For example, children are more likely to be disfluent when producing longer, more syntactically complex sentences (Yaruss, 1999). It is not clear whether this increase is associated with greater demands on the child's language formulation abilities or speech production abilities; however, it is likely that stuttering arises due to the interaction between linguistic and motoric functions. As a group, children who stutter have been shown to exhibit language formulation and speech production abilities that are slightly lower than their typically fluent peers; however, these differences do not generally represent clinically identifiable deficits in speech or language development (Bernstein Ratner, 1997a). Finally, temperament, and specifically the child's sensitivity to stimuli in the environment as well as to speaking mistakes, has been implicated as a factor contributing to the likelihood that a child will react negatively to speech disfluencies (Conture, 2001).

Onset, Development, and Distribution. The onset of childhood stuttering typically occurs between the ages of 2 ½ and 5, though later onset is sometimes reported. Stuttering can develop gradually (with increasing frequency of disfluency and growing severity of individual instances of disfluency), or it can appear relatively suddenly (with the rapid development of more severe stuttering behaviors). Stuttering often begins during a period of otherwise normal or possibly even advanced speech-language development, although many children who stutter exhibit concomitant deficits in other aspects of speech and language development. For example, 30%–40% of preschool children who stutter also exhibit a disorder of speech sound production (articulation or phonological disorder), though the exact nature of the relationship between these communication disorders is not clear (Yaruss, LaSalle, and Conture, 1998).

The lifetime incidence of stuttering may be as high as 5%, although the prevalence is only approximately 1%, suggesting that the majority of young children who stutter—perhaps as many as 75%—recover from stuttering and develop normal speech fluency (Yairi and Ambrose, 1999). Children who recover typically do so within the first several months after onset; however, recovery is also common within the first 2 years after onset (Yairi and Ambrose, 1999). After this time, natural or unaided recovery is less common, and children appear to be significantly less likely to experience a complete recovery if they have been stuttering for longer than 2 to 3 years, or if they are still stuttering after approximately age 7 (Andrews and Harris, 1964). Boys are affected more frequently than girls: in adults, the male to female ratio is approximately 4 or 5 to 1, though at onset, the ratio is closer to 2 to 1, suggesting that girls are more likely to experience recovery than boys. This is particularly true for girls who have other females in their family with a history of recovery from stuttering (Ambrose, Yairi, and Cox, 1997).

Diagnosis and Assessment. The high rate of recovery from early stuttering indicates a positive prognosis for many preschool children who stutter; however, it also complicates the diagnostic process and makes it difficult to evaluate the efficacy of early intervention. There is general agreement among practitioners that it is best to evaluate young children soon after the onset of stuttering to estimate the likelihood of recovery. Often, however, it is difficult to make this determination, and there is considerable disagreement about whether it is best to enroll children in treatment immediately or wait to see whether they will recover without intervention (e.g., Bernstein Ratner, 1997b; Curlee and Yairi, 1997).

Based on the understanding that the etiology of childhood stuttering involves multiple interacting factors, the diagnostic assessment of a preschool child who stutters involves evaluation of several aspects of the child's speech, language, and overall development, as well as selected aspects of the child's environment. Specifically, a complete diagnostic assessment includes the following: (1) a detailed interview with parents or caregivers about factors such as family history of stuttering, the family's reactions to the child's speaking difficulties, the child's reactions to stuttering, the speech and language models the child is exposed to at home, and any other information about communication or other stressors the child may be experiencing, such as competition for talking time with siblings; (2) assessment of the observable characteristics of the child's fluency, including the frequency, duration, and type of disfluencies and a rating of stuttering severity; (3) assessment of the child's speaking abilities, including an assessment of speech sound production/phonological development and oral-motor skills; (4) assessment of the child's receptive and expressive language development, including morphological structures, vocabulary, syntax, and pragmatic interaction; and, increasingly, (5) assessment of the child's temperament, including sensitivity to stimuli in the environment and concerns about speaking difficulties.

Together, these factors can be used to estimate the likelihood that the child will recover from early stuttering without intervention or whether treatment is indicated. Although it is impossible to determine with certainty which children will recover from stuttering without intervention, some diagnostic signs that may indicate an increased risk of continued stuttering and the need for treatment include a family history of chronic stuttering, significant physical tension or struggle during stuttered or fluent speech, concomitant disorders of speech or language, and a high degree of concern about speaking difficulties on the part of the child or the family (e.g., Yairi et al., 1996). Importantly, some practitioners also recommend treatment even in cases where the estimated risk for continued stuttering is low, either in an attempt to speed up the natural recovery process or to
help concerned parents reduce their worries about their child’s fluency.

**Treatment.** As theories about stuttering have changed, so too have preferred treatment approaches, particularly for older children and adults who stutter. At present, there are two primary approaches to treatment for young children who stutter, traditionally labeled “indirect” and “direct” therapy. Indirect therapy is based on the notion that children’s fluency is affected by specific characteristics of their speech, such as speaking rate, time allowed for pausing between words and phrases, and the length and complexity of utterances. Specifically, it appears that children are less likely to stutter if they speak more slowly, allow more time for pausing, or use shorter, simpler utterances. If children can learn to use these “fluency facilitating” strategies, they are more likely to be fluent and, presumably, less likely to develop a chronic stuttering disorder.

A key assumption underlying the indirect treatment approach is that these parameters of children’s speech are influenced by the communication model of the people in the child’s environment. Thus, in indirect therapy, clinicians teach parents and caregivers to use a slower rate of speech, to increase their rate of pausing, and, in some instances, to modify the length and complexity of their utterances, although there is increasing concern among some researchers that restricting the language input children receive may have unintended negative consequence for children’s overall language development (e.g., Bernstein Ratner and Silverman, 2000). Furthermore, although it is clear that children do learn certain aspects of communication from their environment, there is relatively little empirical support for the notion that changing parents’ speech characteristics directly influences children’s speech characteristics or their speech fluency. Even in the absence of clear efficacy data, however, the indirect approach is favored by many clinicians who are hesitant to draw attention to the child’s speech or increase the child’s concerns about their speech fluency.

In recent years, a competing form of treatment for preschool children who stutter has gained popularity (Harrison and Onslow, 1999). This behavioral approach is based on parent-administered intermittent reinforcement of fluent speech and occasional, mild, supportive correction of stuttered speech. Specifically, when a child stutters, the parent labels the stuttering as “bumpy” speech and encourages the child to repeat the sentence “without the bumps.” Efficacy data indicate that this form of treatment is highly successful at reducing the observable characteristics of stuttering, although questions remain about the mechanism responsible for this improved fluency.

Regardless of the approach used in the preschool years, clinicians generally shift from indirect to more direct forms of treatment as children grow older and their awareness of their speaking difficulties increases. Direct treatment strategies include specifically teaching children to use a slower speaking rate or reduced physical tension to smooth out their speech and helping them learn to modify individual instances of stuttering so they are less disruptive to communication (Ramig and Bennett, 1997). Other direct approaches include operant treatments based on reinforcing fluent speech in a hierarchy of utterances of increasing syntactic complexity and length (Bothe, 2002).

A critical component of treatment for older children who stutter, for whom complete recovery is less likely—and even for preschool children who are concerned about their speech or who have significant risk factors indicating a likelihood of continued stuttering (Logan and Yaruss, 1999)—is learning to accept stuttering and to minimize the impact of stuttering in daily activities. As children learn to cope with their stuttering, they are less likely to develop the negative reactions that characterize more advanced stuttering, so the disorder is less likely to become debilitating for them. In addition to pursuing treatment, many older children and families of children who stutter also find meaningful support through self-help groups, and clinicians are increasingly recommending support group participation for their young clients and their families.

**Summary.** Stuttering is a complex communication disorder involving the production of certain types of speech disfluencies, as well as the affective, behavioral, and cognitive reactions that may result. There is no one known cause of stuttering. Instead, childhood stuttering appears to arise because of a complex interaction among several factors that are both genetically and environmentally determined, such as the child’s linguistic abilities, motoric abilities, and temperament. Therefore, diagnostic evaluations of preschoolers who stutter must examine the child’s environment and several aspects of the child’s development. Treatment options include both indirect and direct approaches designed not only to minimize the occurrence of speech disfluencies, but also to minimize the impact of those disfluencies on the lives of children who stutter.

See also LANGUAGE IN CHILDREN WHO STUTTER; STUTTERING.

—J. Scott Yaruss

**References**


Speech Disorders: Genetic Transmission

Advances in behavioral and molecular genetics over the past decade have made it possible to investigate genetic factors that may contribute to speech sound disorders. These speech sound disorders of unknown etiology were often considered to be “functional” or learned. Mounting evidence suggests that at least some speech sound disorders may in part be genetic in origin. However, the search for a genetic basis of speech sound disorder has been complicated by definitional and methodological problems. Issues that are critical to understanding the genetic transmission of speech disorders include prevalence data, phenotype definitions, sex as a risk factor, familial aggregation of disorders, and behavioral and molecular genetic findings.

Prevalence Estimates. Prevalence estimates of speech sound disorders are essential in conducting behavioral and molecular genetic studies. They are used to calculate an individual’s risk of having a disorder as well as to test different genetic models of transmission. Prevalence rates for speech disorders may vary based on the age and sex of the individual, the type of disorder, and the comorbid conditions associated with it. In an epidemiological sample, Shriberg and Austin (1998) reported the prevalence of speech delay in 6-year-old children to be 3.8%. Speech delay was approximately 1.5 times more prevalent in boys than in girls. Shriberg and Austin (1998) also found that children with speech involvement have a two to three times greater risk for expressive language problems than for receptive language problems. Estimates of the comorbidity of receptive language disorders with speech disorders ranged from 6% to 21%, based on whether receptive language was assessed by vocabulary, grammar, or both. Similarly, estimates of comorbidity of expressive language disorders with speech disorders ranged from 38% to 62%, depending on the methods used to assess expressive skills.

Phenotype Definitions. Phenotype definitions (i.e., the behavior that is under study) are also crucial for genetic studies of speech disorders. Phenotype definitions may be broadly or narrowly defined, according to the hypothesis to be tested. A broad phenotype may include language as well as speech disorders and sometimes related language learning difficulties such as reading and spelling disorders (Tallal, Ross, and Curtiss, 1989). An

Further Readings


individual exhibiting a single disorder or a combination of disorders is considered to be affected. Such a broad phenotype may test a general verbal trait deficit hypothesis which holds that there is a common underlying genetic and cognitive basis for speech and language disorders that is expressed differently in individual family members (i.e., variable expression). An alternative explanation is that each disorder has a unique underlying genetic and cognitive basis. Some investigators have narrowly defined the phenotype as a specific speech disorder, such as phonology (Lewis, Ekelman, and Aram, 1989). Even if the proband (i.e., the child with a disorder from whom other family members are identified) is selected by a well-defined criterion, nuclear family members often present a varied spectrum of disorders. Some studies, while narrowly defining the phenotype for the proband, have used a broad phenotype definition for family members. Since older siblings and parents often do not demonstrate speech sound errors in their conversational speech, researchers have relied on historical reports, rather than direct observations of the speech disorder. Figure 1 shows a typical family pedigree of a child with a speech disorder.

Figure 1. A typical family pedigree of a child with a speech disorder. The arrow indicates the proband child. Male family members are represented by squares and female family members are represented by circles. Individuals who are affected with speech disorders are shaded in black. Other disorders are represented by circles. Individuals who are affected family members are represented by squares and female family members.

Sex as a Risk Factor. A robust finding in studies of familial speech and language disorders has been a higher prevalence of disorders in males than in females, ranging from a 2:1 to a 3:1 ratio (Neils and Aram, 1986; Tallal et al., 1989; Tomblin, 1989; Lewis, 1992). Explanations for this increased prevalence in males include referral bias (Shaywitz et al., 1990), immunoreactive theories (Robinson, 1991), differences in rates and patterns of neurological maturation (Plante, 1996), variation in cognitive phenotypes (Bishop, North, and Donlan, 1995), and differences in genetic transmission of the disorders. An X-linked mode of transmission of speech and language disorders has not been supported by pedigree studies (Lewis, 1992; Beitchman, Hood, and Inglis, 1992). However, a sex-specific threshold hypothesis that proposes that girls have a higher threshold for expression of the disorder, and therefore require a higher genetic loading (more risk genes) before the disorder is expressed, has been supported (Tomblin, 1989; Lewis, 1992; Beitchman, Hood, and Inglis, 1992). Consistent with this hypothesis, a higher percentage of affected relatives are reported for female (38%) than for male probands (26%). Differing sex ratios may be found for various subtypes of phonology disorders (Shriberg and Austin, 1998). Hall, Jordan, and Robin (1993) reported a 3:1 male to female ratio for developmental apraxia of speech. Similarly, boys with phonology disorders were found to have a higher rate of comorbid language disorders than girls with phonology disorders (Shriberg and Austin, 1998). Lewis et al. (1999) found that probands with phonology disorders alone demonstrated a more equal sex ratio (59% male and 41% female) than probands with phonology disorders with language disorders (71% male and 29% female).

Familial Aggregation. Familial aggregation refers to the percentage of family members demonstrating a disorder. Familial aggregation or family resemblance may be due to heredity, shared family environment, or both. Research has supported the conclusion that speech and language disorders aggregate within families (Neils and Aram, 1986; Tallal, Ross, and Curtiss, 1989; Tomblin, 1989; Gopnik and Crago, 1991; Lewis, 1992; Felsenfeld, McGuie, and Broen, 1995; Lahey and Edwards, 1995; Spitz et al., 1997; Rice, Haney, and Wexler, 1998). Reports indicate that 23%–40% of first-degree family members of individuals with speech and language disorders are affected. Differences in reported rates of affected family members again may be attributed to differences in definitional criteria for probands and family members.

Two studies specifically examined familial aggregation of phonology disorders (Lewis, Ekelman, and Aram, 1989; Felsenfeld, McGuie, and Broen, 1995). Both studies reported 33% of first-degree family members (nuclear family members) to have had speech-language difficulties. Brothers were most frequently affected.
Behavioral and Molecular Genetic Studies. The twin study paradigm has been employed to identify genetic and environmental contributions to speech and language disorders. Twin studies compare the similarity (concordance) of identical or monozygotic twins to fraternal or dizygotic twins. If monozygotic twins are more concordant than dizygotic twins, a genetic basis is implied. To date, a twin study specifically examining speech sound disorders has not been conducted. Rather, twin studies have employed a broad phenotype definition that includes both speech and language disorders. Twin studies of speech and language disorders (Lewis and Thompson, 1992; Bishop, North, and Donlan, 1995; Tomblin and Buckwalter, 1998) have consistently reported higher concordance rates for monozygotic than for dizygotic twin pairs, confirming a genetic contribution to these disorders. Concordance rates for monozygotic twins range from .70 (Bishop, North, and Donlan, 1995) to .86 (Lewis and Thompson, 1992) and .96 (Tomblin and Buckwalter, 1998). Concordance rates reported for dizygotic twin pairs are as follows: .46 (Bishop, North, and Donlan, 1995), .48 (Lewis and Thompson, 1992), and .69 (Tomblin and Buckwalter, 1998). A large twin study (3000 pairs of twins) suggested that genetic factors may exert more influence at the lower extreme of language abilities, whereas environmental factors may influence normal language abilities more (Dale et al., 1998). These studies, while supporting a genetic contribution to speech and language skills, also indicate a moderate environmental influence. Environmental factors working with genetics may determine speech and language impairment in an individual.

Consistent with these findings, an adoption study by Felsenfeld and Plomin (1997) demonstrated that a history of speech and language disorders in the biological parent best predicted whether or not a child was affected. This relationship was not found when the family history of the adoptive parents was considered. As with twin studies, a broad phenotype definition that encompassed both speech and language disorders was employed. Adoption studies have not been conducted for speech sound disorders alone.

Segregation analyses examine the mode of transmission of the disorder within a family. Segregation analyses have confirmed familial aggregation of speech and language disorders and supported both a major locus model and a polygenic model of transmission of the disorder (Lewis, Cox, and Byard, 1993). The failure to define a definitive mode of transmission may be due to genetic heterogeneity (i.e., more than a single underlying genetic basis).

Only a single study to date has reported a molecular genetic analysis of a family with apraxia of speech and other language impairments. Genetic studies of a single large pedigree, known as the K.E. family, revealed linkage to a region of chromosome 7 (Vargha-Khadem et al., 1995; Fisher et al., 1998). Subsequently, the FOXP2 gene that is postulated to result in the development of abnormal neural structures for speech and language was identified (Lai et al., 2001). Neuroimaging of family members indicated abnormalities in regions of the frontal lobe and associated motor systems. This was the first study that provided direct evidence for a genetic basis for a speech sound disorder associated with a neurological abnormality. It was the initial step in the application of molecular genetic techniques to the study of speech and language disorders. Further studies are needed to determine if FOXP2 is found in other families with speech disorders.

See also SPEECH DISORDERS IN CHILDREN: DESCRIPTIVE LINGUISTIC APPROACHES; DEVELOPMENTAL APRAXIA OF SPEECH; PHONOLOGICAL ERRORS, RESIDUAL.

—Barbara A. Lewis

References


Speech Disorders in Adults, Psychogenic

The human communication system is vulnerable to changes in the individual’s emotional or psychological state. Several studies show the human voice to be a sensitive indicator of different emotions (Aronson, 1991). Psychiatrists routinely evaluate vocal (intensity, pitch), prosodic (e.g., rhythm, rate, pauses), and other features of communication to diagnosis neurotic states (Brodnitz, 1981). Speech-language pathologists also consider the contributions of psychopathology in the evaluation and management of acquired adult speech disorders (Sapir and Aronson, 1990). This is important because depression and/or anxiety are common in stroke, traumatic brain injury, and progressive neurological disease (Giannoni, 1972). Depression frequently occurs after a laryngectomy and may interfere with rehabilitation efforts (Rollin, 1987). Individuals subjected to prolonged stress may speak with excessive tension in the vocal mechanism. When misuse of the vocal apparatus occurs for a long period of time, it can lead to the formation of vocal fold lesions (e.g., nodules) and long-term dysphonia (Aronson, 1991; Case, 1991). One of the responsibilities of the speech-language pathologist is to determine if and how psychogenic components contribute to acquired speech disorders in adults secondary to structural lesions and neurological disease to enhance differential diagnosis and to plan appropriate intervention (Sapir and Aronson, 1990).

In general, some degree of psychopathology is present in most acquired adult speech disorders. It is reasonable that persons who previously communicated normally would be affected psychologically when their ability to communicate was disrupted. In such cases, psychopathology (e.g., depression and anxiety) contributes to and possibly exacerbates the speech disorder, but it is not the cause of the disorder. Therefore, while the speech-language pathologist must be alert to the role of psychopathology in assessment and management of these cases, these disorders are not “purely” psychogenic in nature.

Purely psychogenic speech disorders, the subject of this chapter, are rare in clinical practice. With psychogenic speech disorders, the communication breakdown stems from a conversion disorder. Conversion disorders are included within a larger family of psychiatric disorders, somatoform illnesses. These tend to be associated with pathologic beliefs and attitudes on the part of the patient that results in somatic symptoms. The American Psychiatric Association (1987) defines a conversion disorder as “an alteration or loss of physical functioning that suggests a physical disorder, that actually represents an expression of a psychological conflict or need.” An example might be a woman who suddenly loses her voice because she cannot face the psychological conflict of a spouse’s affair. Here the symptom (voice loss) constitutes a lesser threat to her psychological equilibrium than confronting the husband with his infidelity. A partial list of psychogenic speech disorders includes partial (dysphonia) or complete (aphonia) loss of voice (Andersson and Schalen, 1998), dysarthria (Kallen, Marshall, and Casey, 1986), mutism (Kalman and Granet, 1981), and stuttering (Wallen, 1961; Deal, 1982). There are a few reports that indicate patients can also develop psychogenic language disorders, specifically aphasia (Iddings and Wilson, 1981; Sevush and Brooks, 1983) and dyslexia and dysgraphia (Master and Lishman, 1984). Reports of psychogenic swallowing disorders (dysphagia) also exist (Carstens, 1982).
Diagnosis

Before proceeding with the evaluation and treatment of a patient with a suspected psychogenic speech disorder (PSD), the speech-language pathologist should be sure that the patient has been seen by the appropriate medical specialist to rule out an organic cause of the problem. Usually this is not a problem because the PSD patient will have sought care from several specialists, sometimes simultaneously, and may have a complicated medical history in which multiple diagnoses are tenable (Case, 1991). Therefore, definitive diagnosis of a speech disorder as “psychogenic” is an inexact process relying greatly on the examiner’s skill, experience, and ability to synthesize information from the clinical interview and examination.

Interview and History

The PSD patient often has a history of other conversion disorders and prior psychological stress unrelated to the symptom (Pincus and Tucker, 1974). PSD onset is sudden and tends to be linked to a specific traumatic event (e.g., surgery) or painful emotional experiences (e.g., death of a family member). In the patient interview, the speech-language pathologist determines what the patient might gain from the presenting speech disorder. Primary gain refers to the reduction of anxiety, tension, and conflict provided by the speech disorder. This could be related to a breakdown in communication between the patient and some person of importance, such as a spouse, a boss, or parent. Here the speech problem constitutes a lesser dilemma for the individual than the interpersonal problems from which it arose. Secondary gain refers to those benefits received by the individual from the external environment. This could take the form of monetary compensation, attention, and sympathy from others over perceived distress, being given fewer responsibilities (e.g., work, child care), release from social obligations, and satisfaction of dependency needs. Sometimes secondary gains reinforce the PSD and prolong its duration (Morrison and Rammage, 1994). Unlike individuals with acquired adult communication disorders resulting from structural damage or neurologically disease, the PSD patient may show unusual calmness and lack of concern over the speech disorder, a phenomenon called la belle indifférence. It should be understood, however, that PSD patients do not consciously produce the symptom affecting communication but truly believe they are ill. As a consequence, many PSD patients display excessive concern about their bodies and overreact to normal somatic stimuli.

Clinical Examination

Results of the oral-peripheral and laryngeal examination are often normal or fail to account for the PSD patient’s symptoms. Instrumental measures have limited value in assessment and diagnosis of assessing the individual with a PSD, since these contribute to the patient’s fixation that his or her problem is organically based. The most valuable diagnostic information is gleaned from the clinical examination and the interview. PSD symptoms fluctuate widely across patients and within the same patient. Examples of SDSs affecting voice include conversion aphonia (no voiced but articulated air stream), conversion dysphonia (some voice but abnormal pitch, loudness, or quality), and conversion muteness (no voice but moving of the lips as though articulating) (Case, 1991). Individual PSD patients may report speaking better in some situations than others. The patient may even exhibit variability in symptoms within the context of the clinical examination (Case, 1991). Aphonic or dysphonic patients may vocalize normally when laughing, coughing, or clearing of the throat (Morrison and Rammage, 1994). Further, distraction techniques such as asking the patient to hum, grunt, or say “uh-huh” as an affirmation might produce a normal-sounding voice. PSD patients may react unfavorably when it is pointed out that he or she has produced a normal-sounding voice and insist that the examiner identify a physical cause of the problem.

Intervention

After the examining physician has ruled out an organic cause for the speech disorder and the problem definitively diagnosed as psychogenic, an appropriate intervention plan can be selected. Intervention approaches for the patient with a PSD vary widely and largely depend on the work setting, skill, training, and philosophy of the clinician (see also NEUROGENIC MUTISM, LARYNGECTOMY). For the most part, interventions used by speech-language pathologists focus on removal of the symptoms (e.g., aphonia, dysphonia, dysarthria) found to be abnormal in the evaluation, and give limited attention to psychological or psychiatric issues.

Several studies report results of successful interventions for PSD patients (Aronson, 1969; Marshall and Watts, 1975; Kalman and Granet, 1981; Carstens, 1982; Kallen, Marshall, and Casey, 1986; Andersson and Schalen, 1998). These usually begin with the clinician acknowledging the patient’s distress and assuring the patient that there is no known organic reason for the problem. Potential factors that might contribute to the patient’s distress are discussed in a nonthreatening manner. Here the focus of attention is on how the symptom or symptoms are disrupting communication. Individually designed intervention programs are then initiated and typically proceed in small steps or behavioral increments. For example, a sequence of intervention steps for a patient with psychogenic aphonia may include the following: (1) eliciting a normal vocal tone with a cough, grunt, or hum; (2) prolonging the normal vocalization into a vowel; (3) turning the vowel into a VC word; (4) linking two VC words together; (5) producing sentences using several VC words; and so forth. Relaxation exercises, patient education, and counseling may be included in the intervention programs for some PSD patients. In general, symptomatic interventions are successful with two to 10 sessions of treatment.
Three factors—duration of time elapsing between onset of symptoms and beginning of therapy, severity of the speech disorder, and the degree to which the patient's speech disorder is distinguishable from the emotional disturbance—are related to intervention success (Case, 1991). Generally, intervention is more apt to be successful if the time between the onset of symptoms and the start of intervention is short, if the problem is severe, and if a speech disorder is clearly distinguishable from the emotional disturbance responsible for the problem (Case, 1991). However, certain patients may be more anxious for help and receptive to therapy if they have been without voice for a considerable period of time before consulting a speech-language pathologist (Freeman, 1986). In addition, patients who react unfavorably to intervention success or refuse to accept or acknowledge their "improved communicative status" do not respond well to behavioral treatments and have a poor prognosis for symptom-based interventions. In such cases the speech-language pathologist must refer the patient to the psychologist or the psychiatrist.

—Robert C. Marshall

References


Further Readings


Speech Disorders in Children: A Psycholinguistic Perspective

The terminology used to describe speech problems is rooted in classificatory systems derived from different academic disciplines. In order to understand the rationale behind the psycholinguistic approach, it is helpful to examine other approaches and compare how speech problems have been classified from different perspectives. Three perspectives that have been particularly influential are the medical, linguistic, and psycholinguistic perspectives.

In a medical perspective, speech and language problems are classified according to clinical entity. Commonly used labels include dyspraxia, dysarthria, and stuttering. Causes of speech difficulties can be identified (e.g., cleft palate, hearing loss, neurological impairment) or an associated medical condition is known (e.g., autism, learning difficulties, Down syndrome).

Viewing speech and language disorders from a medical perspective can be helpful in various ways. First, through the medical exercise of constructing a differential diagnosis, a condition may be defined when symptoms commonly associated with that condition are identified; two examples are dyspraxia and dysarthria. Second, for some conditions, medical management can contribute significantly to the prevention or remediation of the speech or language difficulty, such as by insertion of a cochlear implant to remediate hearing loss or by surgical repair of a cleft palate. Third, the medical perspective may be helpful when considering the prognosis for a child’s speech and language development, such as when a progressive neurological condition is present.

However, the medical approach has major limitations as a basis for the principled remediation of speech problems in individual children. A medical diagnosis cannot always be made. More often the term “specific speech and/or language impairment” is used once all other possible medical labels have been ruled out. Moreover, even if a neuroanatomical correlate or genetic basis for a speech and language impairment can be identified, the medical diagnosis does not predict with any precision the speech and language difficulties that an individual child will experience, so the diagnosis will not significantly affect the details of a day-to-day intervention program. To plan appropriate therapy, the medical model needs to be supplemented by a linguistic approach.

The linguistic perspective is primarily concerned with the description of language behavior at different levels of analysis. If a child is said to have a phonetic or articulatory difficulty, the implication is that the child has problems with the production of speech sounds. A phonological difficulty refers to inability to use sounds contrastively to convey meaning. For example, a child may use [t] for [s] at the beginning of words, even though the child can produce a [s] sound in isolation perfectly well. Thus, the child fails to distinguish between target words (e.g., “sea” versus “tea”) and is likely to be misunderstood by the listener. The cause of this difficulty may not be obvious.

The linguistic sciences have provided an indispensable foundation for the assessment of speech and language difficulties (Ingram, 1976; Grunwell, 1987). However, this assessment is still a description and not an explanation of the disorder. Specifically, a linguistic analysis focuses on the child’s speech output but does not take account of underlying cognitive processes. For this, a psycholinguistic approach is needed.

The psycholinguistic approach attempts to make good some of the shortcomings of the other approaches by viewing children’s speech problems as being derived from a breakdown in an underlying speech processing system. This assumes that the child receives information of different kinds (auditory, visual) about an utterance, remembers it, and stores it in a variety of lexical representations (a means for keeping information about words, which may be semantic, grammatical, phonological, motor, or orthographic) within the lexicon (a store of words), then selects and produces spoken and written words. Figure 1 illustrates the basic essentials of a psycholinguistic model of speech processing. On the left there is a channel for the input of information via the ear and on the right a channel for the output of information through the mouth. The lexical representations at the top of the model store previously processed information. In psycholinguistic terms, top-down processing refers to an activity whereby previously stored information (i.e., in the lexical representations) is helpful and used, for example, in naming objects in pictures. A bottom-up processing activity requires no such prior knowledge and can be completed without accessing stored linguistic knowledge from the lexical representations; an example is repeating sounds.

A number of models have been developed from this basic structure (e.g., Dodd, 1995; Stackhouse and Wells, 1997; Hewlett, Gibbon, and Cohen-McKenzie, 1998; Chiat, 2000). Although these models differ in their presentation, they share the premise that children’s speech
difficulties arise from one or more points in a faulty speech processing system. The aim of the psycholinguistic approach is to find out exactly where a child’s speech processing skills are breaking down and how these deficits might be compensated for by coexisting strengths. The investigative procedure to do so entails generating hypotheses, normally from linguistic data, about the levels of breakdown that give rise to disordered speech output. These hypotheses are then tested systematically through carefully constructed tasks that provide sufficient data to assemble a child’s profile of speech processing strengths and weaknesses (Stackhouse and Wells, 1997; Chiat, 2000).

Collation of these profiles shows that some children with speech difficulties have problems only on the output side of the model. However, many children with persisting speech problems have pervasive speech processing difficulties (in input, output, and lexical representations) that impede progress. For example, when rehearsing new words for speech or spelling, it is usual to repeat them verbally. An inconsistent or distorted output, normally the result of more than one level of breakdown, may in turn affect auditory processing skills, memory, and the developing lexicon. It is therefore not surprising that children with dyspraxic speech difficulties often have associated input (Bridgeman and Snowling, 1988) and spelling difficulties (Clarke-Klein and Hodson, 1995; McCormick, 1995).

The case study research of children with developmental speech disorders, typical of the psycholinguistic approach, has shown that not only are children unintelligible for different reasons but also that different facets of unintelligibility in an individual child can be related to different underlying processing deficits (Chiat, 1983, 1989; Stackhouse and Wells, 1993). Extending the psycholinguistic approach to word finding difficulties, Constable (2001) has discovered that such difficulties are long-term consequences of underlying speech processing problems that affect how the lexical representations are stored, and in particular how the phonological, semantic, and motor representations are interconnected.

The psycholinguistic approach has also been used to investigate the relationship between spoken and written language and to predict which children may have long-term difficulties (Dodd, 1995; Stackhouse, 2001). Those children who fail to progress to a level of consistent speech output, age-appropriate phonological awareness, and letter knowledge skills are at risk for literacy problems, particularly when spelling. Psycholinguistic analysis of popular phonological awareness tasks (e.g., rhyme, syllable/sound segmentation and completion, blending, spoonerisms) has shown that the development of phonological awareness skills depends on an intact speech processing system (Stackhouse and Wells, 1997). Thus, children with speech difficulties are disadvantaged in school, since developing phonological awareness is a necessary stage in dealing with alphabetic scripts such as English. Further, these phonological awareness skills are needed not just for an isolated activity, such as a rhyme game, but also to participate in the interactions typical of phonological intervention sessions delivered by a teacher or clinician (Stackhouse et al., 2002).

An individual child’s psycholinguistic profile of speech processing skills provides an important basis for planning a targeted remediation program (Stackhouse and Wells, 2001). There is no prescription for delivering this program, nor is there a bag of special activities. All intervention materials have the potential to be used in a psycholinguistic way if analyzed appropriately. Principled intervention is based on setting clear aims (Rees, 2001a). Tasks are chosen or designed for their psycholinguistic properties and manipulated to ensure appropriate targeting and monitoring of intervention. To this end, each task is analyzed into its components, as follows:

\[
\text{Task} = \text{Materials} + \text{Procedure} + \text{Feedback} \pm \text{Technique}
\]

Rees (2001b) presents seven questions for examining these four components. She demonstrates how altering any one of them can change the nature of the task and thus the psycholinguistic demands made on the child.

In summary, a psycholinguistic approach to intervention puts the emphasis first on the rationale behind the design and selection of tasks for a particular child, and then on the order in which the tasks are to be presented to a child so that strengths are exploited and weaknesses supported (Vance, 1997; Corrin, 2001a, 2001b; Waters, 2001). The approach tackles the issues of what to do, with whom, why, when, and how.

In a review of psycholinguistic models of speech development, Baker et al. (2001) present both box-
and-arrow and connectionist models as new ways of conceptualizing speech impairment in children. This discussion has focused on the former, since to date, box-and-arrow models have arguably had the most impact on clinical practice by adding to our repertoire of assessment and treatment approaches, and also by promoting communication and collaboration between teachers and clinicians (Popple and Wellington, 2001). The success of the psycholinguistic approach may lie in the fact that it targets the underlying sources of difficulties rather than the symptoms alone (Holm and Dodd, 1999). Although it is true that the outcome of intervention depends on more than a child's speech processing profile (Goldstein and Geirut, 1998), the development of targeted therapy through the setting of realistic aims and quantifiable objectives should make a contribution to the measurement of the efficacy of intervention.

—Joy Stackhouse

References


Further Readings


Constable, A., Stackhouse, J., and Wells, B. (1997). Developmental word-finding difficulties and phonological processing:
Speech Disorders in Children: Behavioral Approaches to Remediation

Speech disorders in children include articulation and phonological disorders, stuttering, cluttering, developmental apraxia of speech, and a variety of disorders associated with organic conditions such as brain injury (including cerebral palsy), cleft palate, and genetic syndromes. Despite obvious linguistic influences on the analysis, classification, and theoretical understanding of speech disorders in children, most current treatment methods use behavioral techniques. The effectiveness of behavioral treatment techniques in remediating speech disorders in children has been well documented (Onslow, 1993; Bernthal and Bankson, 1998; Hegde, 1998, 2001; Pena-Brooks and Hegde, 2000). Behavioral techniques that apply to all speech discords—and indeed to most disorders of communication—include positive reinforcement and reinforcement schedules, negative reinforcement, instructions, demonstrations, modeling, shaping, prompting, fading, corrective feedback to reduce undesirable responses, and techniques to promote generalized productions and response maintenance.

A basic procedure in implementing behavioral intervention is establishing the baserates of target behaviors. Baserates, or baselines, are systematically measured values of specified behaviors or skills in the absence of planned intervention. Baserates are the natural rates of response when nothing special (such as modeling or explicit positive reinforcement) is programmed. Baserates help establish a stable and reliable response rate against which the effects of a planned intervention or an experimental treatment can be evaluated. The baserate of any parameter should be determined by at least three measurements to establish their stability. For instance, to establish the baserates of stuttering in a child, the clinician should measure stuttering in at least three consecutive speech samples. Baserates also should sample responses adequately. For instance, to establish the baserate of production of a phoneme in a child, 15–20 words, phrases, or sentences that contain the target phoneme should be used. Baserates also may be established for different settings, such as the clinic, classroom, and home. In each setting, multiple measurements would be made.

Positive Reinforcement and Reinforcement Schedules. Positive reinforcement is a powerful method of shaping new behaviors or increasing the frequency of low-frequency but desired behaviors. It is a method of selecting and strengthening an individual’s behaviors by arranging for certain consequences to occur immediately after the behavior (Skinner, 1953, 1969, 1974). In using positive reinforcement, the clinician arranges a behavioral contingency, which is an interdependent relationship between a response made in the context of a stimulus array and the consequence that immediately follows it. Therefore, technically, behavioral contingency is the heart of behavioral treatment.

Positive reinforcers are specific events or objects that, following a behavior, increase the future probability of that behavior. Speech-language pathologists routinely use a variety of positive reinforcers in teaching speech skills to children and adults. Praise is a common positive reinforcer. Other positive reinforcers include tokens, given for correct responses, that may be exchanged for small gifts. Biofeedback or computer feedback as to the accuracy of response are other forms of positive reinforcement.

Positive reinforcers are initially offered for every correct response, resulting in a continuous reinforcement schedule. When the new response or skill has somewhat stabilized, the reinforcer may be offered for every nth response, resulting in a fixed ratio (FR) schedule. For instance, a child may receive reinforcer for every fifth correct phoneme production (an FR5). Gradually reducing the number of reinforcers with the use of progressively larger ratios will help maintain a skill taught in clinical settings.

Antecedent Control of Target Behaviors. A standard behavioral method is to carefully set the stage for a skill to be taught in treatment sessions. This technique, known as antecedent control, increases the likelihood of a target response by providing stimuli that evoke it. Stimulus manipulations include a variety of procedures, such as modeling, shaping, prompting, and fading.

Modeling. In modeling, the clinician produces the target response, which is then expected to be followed by at least an attempt to produce the same response by the client. The clinician’s behavior is the model, which the client attempts to imitate (response). In most cases, modeling is preceded by instructions to the client on
how to produce a response, and demonstrations of target responses. Although instructions and demonstrations are part of behavioral treatment procedures, much formal research has focused on modeling as a special stimulus to help establish a new target response.

A child’s initial attempt to imitate a target response modeled by the clinician may be more or less correct; nonetheless, the clinician might wish to reinforce all attempts in the right direction. In gradual steps, the clinician may then require responses that are more like the modeled response. To achieve this final result of an imitated response that matches the modeled stimulus, shaping is often used.

**Shaping.** Whereas straightforward positive reinforcement is effective in increasing the frequency of a low-frequency response, shaping is necessary to create skills that are absent. Shaping, also known as successive approximations, is a procedure to teach new responses in gradual steps. The entire procedure typically includes instructions, demonstrations, modeling, and positive reinforcement. A crucial aspect of shaping is specifying the individual components of a complex response and teaching the components sequentially in a manner that will result in the final target response. For instance, in teaching the correct production of /s/ to a child who has an articulation disorder, the clinician may identify such simplified components of the response as raising the tongue tip to the alveolar ridge, creating a groove along the tongue tip, approximating the two dental arches, blowing air through the tongue-tip groove, and so forth. The child’s production of each component response is positively reinforced and practiced several times. Finally, the components are put together to produce the approximation of /s/. Subsequently, and in progressive steps, better approximations of the modeled sound production are reinforced, resulting in an acceptable form of the target response. To further strengthen a newly learned response, the clinician may use the prompting procedure.

**Prompting.** The probability of a target response that has just emerged with assistance from the previously described procedures may fluctuate from moment to moment. The child may appear unsure and the response rate may be inconsistent. In such cases, prompting will help stabilize that response and increase its frequency. Prompting is a special cue or a stimulus that will help evoke a response from an unsure client. Such cues take various verbal and nonverbal forms. Examples of verbal prompts include such statements as “What do you say to this picture?” “The word starts with a /p/” (both prompt a correct naming or articulatory response). Nonverbal prompts include a variety of facial and hand gestures that suggest a particular target response; the meaning of some gestures may first have to be taught to the child. For instance, a clinician might tell a child who stutters that his or her speech should be slowed down when a particular hand gesture is made. In prompting the production of a phone such as /p/, the clinician may press the two lips together, which may lead to the correct production. Eventually, the influence of such special stimuli as prompts is reduced by a technique called fading.

**Fading.** Fading is a technique used to gradually withdraw a special stimulus, such as models or prompts, while still maintaining the correct response rate. Abrupt withdrawal of a controlling stimulus will result in failure to respond. In fading, a modeled stimulus or prompt may be reduced in various ways. For instance, a modeling such as “Say I see sun” (in training the production of phoneme /s/) may be shortened by the clinician to “Say I see...”; or the vocal intensity of the prompter or modeler may be reduced such that the modeling becomes progressively softer and eventually inaudible, with only articulatory movements (e.g., correct tongue position) being shown.

**Corrective Feedback.** Various forms of corrective feedback may be provided to reduce the frequency of incorrect responses. Verbal feedback such as “that is not correct,” “that was bumpy speech” (to correct stuttering in young children), “that was too fast,” and so forth is part of all behavioral treatment programs. Additional corrective procedures include token loss for incorrect responses (tokens earned for correct responses and lost for incorrect ones), and time-out, which includes a brief period of no interaction made contingent on incorrect responses. For instance, every stuttering may be followed by a signal to stop talking for 5 seconds.

**Generalization and Maintenance.** Behavioral techniques to promote generalized production of speech skills in natural environments and maintenance of those skills over time are important in all clinical work. Teaching clients to self-monitor the production of their newly acquired skills and training significant others to prompt and reinforce those skills at home are among the most effective of the generalization and maintenance techniques.

—M. N. Hegde

**References**


Further Readings


Speech Disorders in Children: Birth-Related Risk Factors

The youngest clients who receive services from speech-language pathologists are neonates with medical needs. This area of care came into being during the last several decades as the survival rate of infants with medical needs improved and it became apparent that developmental delays would likely be prevalent among the survivors. Since 1970, for example, the survival rate of infants in some low birth weight categories jumped from 30% to 75%; among the survivors the occurrence of mental retardation is 22%–24% (Bernbaum and Hoffman-Williamson, 1991).

The prevalence of developmental delay among previously medically needy neonates resulted in federal and state laws that give these children legal rights to developmental services (Kern, Delaney, and Taylor, 1996). These laws identify the physical and mental conditions and the biological and environmental factors present at birth that are most likely to result in future developmental delay. The purpose in making this identification is to permit an infant to receive intervention services early in life, when brain development is most active and before the negative social consequences of having a developmental delay, including one in the speech domain, can occur (Bleile and Miller, 1994).

The most important federal legal foundation for developmental services for young children with medical needs is the Individuals with Disabilities Education Act (IDEA). This law gives individual states the authority to determine which conditions and factors present at birth place a child at sufficient risk for future developmental delay that the child qualifies for education services. Examples of conditions and factors that the act indicates place an infant at risk for future developmental delay are listed in Table 1. A single child might have several conditions and risk factors. For example, a child born with Down syndrome might also experience respiratory distress as a consequence of the chromosomal abnormality, as well as an unrelated congenital infection.

Many populations of newborn children with medical needs are at high risk for future speech disorders. In part this is because developmental speech disorders are common among all children (Slater, 1992). However, the medical condition or factor itself may contribute to the child having a speech disorder, as occurs with children with cerebral palsy, a tracheotomy, or cleft palate (Bleile, 1993), and the combination of illness and long-term stays in the hospital may limit opportunities for learning, including in the speech domain.

Children with hearing impairment and those with Down syndrome are two relatively large populations with birth-related conditions and factors that are likely to experience future speech disorders (see MENTAL RETARDATION AND SPEECH IN CHILDREN). Two additional relatively large populations of newborn children likely to experience future speech disorders are those born underweight and those whose mother engaged in substance abuse during pregnancy.

In the United States, approximately 8.5% of infants are born underweight (Guyer et al., 1995). The major birth weight categories are low birth weight, very low birth weight, extremely low birth weight, and micro-premie. Birth weight categories as measured in grams and pounds are shown in Table 2. As the category of micro-premie suggests, many low birth weight children are born prematurely. A typical pregnancy lasts 40 weeks from first day of the last normal menstrual cycle; a preterm birth is defined as one occurring before the completion of 37 weeks of gestation. The co-occurrence of low birth weight and prematurity varies by country; in the United States, 70% of low birth weight babies are also born prematurely.

Approximately 36%–41% of women in the United States abuse illicit drugs, alcohol, or nicotine sometime during pregnancy (Center on Addiction and Substance Abuse, 1996). Illicit drugs account for 11% of this

Table 1. Conditions and Factors Present at Birth or Shortly Thereafter with a High Probability of Resulting in Future Developmental Delay

<table>
<thead>
<tr>
<th>Condition or Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosomal abnormalities such as Down syndrome</td>
<td></td>
</tr>
<tr>
<td>Genetic or congenital disorders</td>
<td></td>
</tr>
<tr>
<td>Severe sensory impairments, including hearing and vision</td>
<td></td>
</tr>
<tr>
<td>Inborn errors of metabolism</td>
<td></td>
</tr>
<tr>
<td>Disorders reflecting disturbance of the development of the nervous system</td>
<td></td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td></td>
</tr>
<tr>
<td>Hyperbilirubinemia at levels exceeding the need for exchange transfusion</td>
<td></td>
</tr>
<tr>
<td>Major congenital anomalies</td>
<td></td>
</tr>
<tr>
<td>Congenital infections</td>
<td></td>
</tr>
<tr>
<td>Disorders secondary to exposure to toxic substances, including fetal alcohol syndrome</td>
<td></td>
</tr>
<tr>
<td>Low birth weight</td>
<td></td>
</tr>
<tr>
<td>Respiratory distress</td>
<td></td>
</tr>
<tr>
<td>Lack of oxygen</td>
<td></td>
</tr>
<tr>
<td>Brain hemorrhage</td>
<td></td>
</tr>
<tr>
<td>Nutritional deprivation</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Birth Weight Categories, in Grams and Pounds

<table>
<thead>
<tr>
<th>Categories</th>
<th>Grams</th>
<th>Pounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low birth weight</td>
<td>&lt;2,500</td>
<td>5.5</td>
</tr>
<tr>
<td>Very low birth weight</td>
<td>&lt;1,500</td>
<td>3.3</td>
</tr>
<tr>
<td>Extremely low birth weight</td>
<td>&lt;1,000</td>
<td>2.2</td>
</tr>
<tr>
<td>Micropremies</td>
<td>&lt;800</td>
<td>1.76</td>
</tr>
</tbody>
</table>
substance abuse, and heavy use of alcohol or nicotine accounts for the other 25%–30%. Approximately three-quarters of pregnant women who abuse one substance also abuse other substances (Center on Addiction and Substance Abuse, 1996). For example, a pregnant woman who abuses cocaine might also drink heavily. When alcohol is an abused substance, more severely affected children are considered to have fetal alcohol syndrome. The hallmarks of fetal alcohol syndrome are mental retardation and physical deformities (Streissguth, 1997). Children with milder cognitive impairments and without physical deformities are considered to have fetal alcohol effect or alcohol-related neurodevelopmental disorder.

Regardless of the specific cause of the disorder, speech-disordered clients with birth-related conditions and factors receive services similar to those given other children. The clinician’s primary responsibility is to provide evaluation and intervention services appropriate to the child’s developmental abilities. A difference in care provision is that the child’s speech disorder is likely to occur as part of a larger picture of medical problems and developmental delay. This may make it difficult to diagnose and treat the speech disorder, especially when the child is younger and medical problems may predominate. In addition to having thorough training in typical speech and language development, a speech-language clinician working with these children should possess the following:

• Basic knowledge of medical concepts and terminology
• Ability to access and understand information about unfamiliar conditions and factors as need arises
• Knowledge of safety procedures and health precautions
• Ability to work well with teams that include the child’s caregivers and professionals

A neonate identified as at risk for future developmental delay typically first receives developmental services in a hospital intensive care unit. Medical and developmental services are provided by a team of health care professionals. Often, a primary role of the speech-language clinician is to assess the oral mechanism to determine readiness to feed. Such evaluations are particularly important for clients at risk for aspiration. These include children with neurological and physical handicaps, as well as those born prematurely, whose immature systems of neurological control often do not allow orally presented food to be managed safely. The speech-language pathologist may also counsel the child’s caregivers and offer suggestions about ways to facilitate communication development.

An early intervention program is initiated shortly after the child is born and the risk factor has been identified. The exception is a child born prematurely, whose nervous system may not yet be able to manage environmental stimulation. Such a child typically receives minimal stimulation until the time he or she would have been born if the pregnancy had been full term. Early intervention typically includes a package of services, including occupational and physical therapy, social services, and speech-language pathology. The role of the speech-language pathologist includes assessing communication development and implementing an early intervention program to facilitate the child’s communication abilities. Interacting with the child’s caregivers assumes increasing importance as medical issues resolve, allowing the family to give greater attention to developmental concerns.

The vast majority of children born with medical needs grow to possess the cognitive and physical capacity for speech. For higher functioning children, the clinician’s role includes providing the evaluation and treatment services to facilitate speech and language development. The speech disorders of many high-functioning children resolve by the end of the preschool years or during the early grade school years. Such children are at risk for future reading problems and other learning difficulties, and their progress in communication should continue to be monitored even after the speech disorder has resolved. Speech may prove challenging for children with more extensive developmental problems. A general clinical rule of thumb is that a child who will speak typically will do so by 5 years of age (Bleile, 1995). Children with more limited speech potential may be taught to communicate through a combination of speech and nonverbal options. Lower functioning children might be taught to communicate through an alternative communication system (see AUGMENTATIVE AND ALTERNATIVE COMMUNICATION APPROACHES IN CHILDREN). Some useful web sites for further information include www.asha.org; www.autism.org; www.cdc.gov; www.intelehealth.com; www.mayohealth.org; www.med.harvard.edu; www.modimes.org; www.ncbi.nlm.nih.gov/PubMed/; www.ndss.org; and www.nih.gov.

—Ken Bleile and Angela Burda

References
Further Readings

Speech Disorders in Children: Cross-Linguistic Data

Since the 1960s, the term articulation disorder has been replaced in many circles by the term phonological disorder. This shift has been driven by the recognition that children with articulation disorders show general patterns in their speech that are not easily identified by an articulatory defect approach. Further, these patterns appear similar to those used by younger, typically developing children. The notion of phonological versus articulatory impairment, however, has not been examined in depth through comparisons of typically developing children and children with phonological impairment across a range of languages. Such comparisons would provide the most revealing evidence in support of one view over the other. If articulatory factors are behind children’s phonological impairment, children with such impairments should show patterns somewhat independent of their linguistic environment. They should look more like one another than like their linguistically matched peers. If there is a linguistic basis to phonological impairment, however, it has not been examined in depth through comparisons of typically developing children and children with phonological impairment across a range of languages. Such comparisons would provide more revealing evidence in support of one view over the other. If articulatory factors are behind children’s phonological impairment, children with such impairments should show patterns somewhat independent of their linguistic environment. They should look more like one another than like their linguistically matched peers.

Cross-Linguistic Predictions

Articulatory Deficit

| English TD + Italian TD = different |
| English PI + Italian PI = same |
| English TD + English PI = different |
| Italian TD + Italian PI = different |

Linguistic Deficit

| English TD + Italian TD = different |
| English PI + Italian PI = different |
| English TD + English PI = same |
| Italian TD + Italian PI = same |

The pursuit of this line of research requires two steps. First, it needs to be verified that there are indeed cross-linguistic differences in phonological acquisition between typically developing children. The truth of this claim is not self-evident. Locke (1983), for example, has argued that children will be similar cross-linguistically until some point after the acquisition of the first 50 words. Elsewhere, however, it has been argued that such cross-linguistic differences are evident at the very earliest stages of phonological development (Ingram, 1989). Resolution of this issue will require extensive research into early typical phonological development in a range of languages. To date, the data support early cross-linguistic differences.

The second, critical step is to determine whether children with phonological delay look like their typically developing peers or like children with phonological impairment in other linguistic communities. The data on this issue are even more sparse than for the first step, but some preliminary data exist, and those data support the linguistic account of phonological delay over the articulatory one.

These two steps can be summarized as follows:

Step 1: Verify that typically developing children vary cross-linguistically.

Step 2: Determine whether children with a phonological impairment look like their typically developing peers or like children with phonological impairments in other linguistic communities.

Examination of a range of studies on early phonological development shows that children in different linguistic environments converge in their acquisition toward a basic or core phonetic inventory of speech sounds that is different for each language. This is demonstrated here by examining the production of word-initial consonants in English, French, K’iche’, and Dutch.

Below is an inventory of the English consonants typically used in the early stage of phonological acquisition (Ingram, 1981). English-speaking children show early acquisition of three place features, a voicing contrast among stops, and a series of voiceless consonants.

**English**

| m | n |
| b | d | g |
| p | t | k |
| f | s | h |
| w |

French shows some similarities to English, but also two striking differences, based on my analysis of selected diary studies. French-speaking children tend to acquire velar consonants later, yet show an early use of /j/, a sound that appears later in English.

**French**

| m |
| b | d |
| p | t |
| f | s |
| l |
K’iche’ (formerly spelled Quichè) is a Mayan language spoken in Guatemala. K’iche’ children (Pye, Ingram, and List, 1987) show an early /l/, as in French, and an affricate, /tʃ/, which is one of the most frequent early sounds acquired. The first fricative tends to be the velar /x/, despite the fact that the language has both /s/ and /ʃ/.

K’iche’

<table>
<thead>
<tr>
<th>m n</th>
<th>p t tf k</th>
</tr>
</thead>
<tbody>
<tr>
<td>w l</td>
<td>x j</td>
</tr>
</tbody>
</table>

Lastly, below is an intermediate stage in Dutch, based on data in Beers (1994). The Dutch inventory shows early use of the velar fricative /x/, as in K’iche’, but also the full range of Dutch fricatives, which appear at more or less the same time.

Dutch

<table>
<thead>
<tr>
<th>m n</th>
<th>b d</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k</td>
<td>f s x h</td>
</tr>
<tr>
<td>w j</td>
<td>j</td>
</tr>
</tbody>
</table>

Data such as these from English, French, K’iche’, and Dutch show the widely different ways that children may acquire their early consonantal inventories. It has been proposed that these differences result from the varying roles these consonants play in the phonologies of the languages discussed (Ingram, 1989). The more frequently a consonant is used in a wide range of words, the more likely it is that it will be in the early inventory.

The question now is, what do the inventories of children with phonological impairment look like when comparisons like those above are done? The limited evidence to date indicates that they look like their same language typically developing peers. Data from Italian, Turkish, and Swedish suggest that this is the case.

The first language examined here is Italian, using the data reported in Bortolini, Ingram, and Dykstra (1993). Typically developing Italian children use the affricate /tʃ/ and the fricative /v/, both later acquisitions for English-speaking children. Italian-speaking children with phonological impairment have an inventory that is a subset of the one used by typically developing children. They do not have the affricate but do show the early acquisition of /v/. The early use of /v/ can be traced back to the fact that the voiced labiodental fricative is a much more common sound in the vocabulary of Italian-speaking children than it is in the vocabulary of their English-speaking peers.

Italian

<table>
<thead>
<tr>
<th>Typically Developing</th>
<th>Phonologically Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t tf k</td>
<td>p t k</td>
</tr>
<tr>
<td>b d g</td>
<td>b d</td>
</tr>
<tr>
<td>f s</td>
<td>f s</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
</tr>
</tbody>
</table>

Topbas (1992) provides data on Turkish for both typically developing children and children with phonological impairment. The Turkish inventory of the typically developing children is noteworthy for its lack of early fricatives, despite a system of eight fricatives, both voiced and voiceless, in the adult language. For English-speaking children, the lack of fricatives is often a sign of phonological impairment, but this appears to be expected for typically developing Turkish-speaking children. The data on phonological impairment are from a single child at age 6 years, 0 months. This child shows the same lack of fricatives and the early affricate, just as in the typically developing data.

Turkish

<table>
<thead>
<tr>
<th>Typically Developing</th>
<th>Phonologically Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>m n</td>
<td>m n</td>
</tr>
<tr>
<td>b d</td>
<td>b d</td>
</tr>
<tr>
<td>p t tf k</td>
<td>j j</td>
</tr>
</tbody>
</table>

Lastly, Swedish data are available to pursue this issue further (Magnussen, 1983; Nettelbladt, 1983). The data on typically developing Swedish children are based on a case study of a child at age 2 years, 2 months. This child lacked the velar stop and the voiced fricative /v/. The voiced stops were also missing in a group of ten children with phonological impairment; however, these children did show early use of /v/, just as the Italian-speaking children did.

Swedish

<table>
<thead>
<tr>
<th>Typically Developing</th>
<th>Phonologically Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>m n</td>
<td>m n</td>
</tr>
<tr>
<td>b d</td>
<td>b d</td>
</tr>
<tr>
<td>p t</td>
<td>p t</td>
</tr>
<tr>
<td>f s h</td>
<td>f s</td>
</tr>
<tr>
<td>v j</td>
<td>v j</td>
</tr>
</tbody>
</table>

Thus, preliminary data from a range of languages support the phonological rather than the articulatory account of phonological impairment. This results in two preliminary conclusions: (1) Typically developing children show early phonological inventories unique to their linguistic environment. (2) Children with phonological impairment show systems more similar to those of their typically developing peers in their own linguistic environment than to those of children with phonological impairment in other language environments.

—David Ingram

References


Symposium on Research in Child Language Disorders, University of Wisconsin, Madison.

Speech Disorders in Children: Descriptive Linguistic Approaches

Linguistic approaches to treating children’s speech disorders are motivated by the fact that a phonology is a communication system (Stoel-Gammon and Dunn, 1985; Ingram, 1990; Grunwell, 1997). Within this system, patterns are detectable within and among various subcomponents: (1) syllable and word shapes (phonotactic repertoire), (2) speech sounds (phonetic repertoire), (3) the manner in which sounds contrast with each other (phonemic repertoire), and (4) the behaviors of different sounds in different contexts (phonological processes). Each of these subcomponents may influence the others; each may interfere with successful communication. Therefore, individual sound or structure errors are treated within the context of the child’s whole phonological system rather than one by one; remediation begins at the level of communicative function—the word.

The assumption underlying this type of treatment is that the child’s phonological system—his or her (subconscious) mental organization of the sounds of the language—is not developing in the appropriate manner for the child’s language or at a rate appropriate for the child’s age. The goal is for the child to adjust his or her phonological system in the needed direction. Initiating a change in one part of the phonological system is expected to have a more general impact on the whole system.

The sounds of a language are organizable into various categories according to their articulatory or acoustic features. The consonants /p, t, k, b, d, g/, for example, are all members of the category stops. They are non-continuant because the airflow is discontinued in the oral cavity during their production. Some sounds also have the feature voiceless because they are produced without glottal vibration. Each speech sound can be categorized in different ways according to such features. These features are called distinctive because they differentiate each sound from all others in the language. The (implicit) knowledge of distinctive features is presumed to be one organizational basis for phonological systems.

Distinctive features therapy (McReynolds and Engmann, 1975) focuses on features that a child’s system lacks. A child who produces no fricatives lacks the continuant feature, so distinctive features therapy would focus on this feature. In theory, establishing a continuant in any place of articulation (e.g., [s]) would lead to the child’s generalization of the feature to other, untrained fricatives. Once the feature is included in the child’s feature inventory, it can be combined with other features (e.g., voice, palatal) to yield the remaining English fricatives.

The function of distinctive features is to provide communicative contrast. The more features we use, the more lexical distinctions we make and the more meanings we express. A child who produces several target phonemes identically (e.g., all fricatives as [f]) will have too many homonyms. Therapy will focus on using more distinctive features, for example, producing different target phonemes distinctively. Williams (2000a, 2000b) recommends simultaneously contrasting all target sounds with the overgeneralized phoneme. In some disordered phonologies, features are noncontrastive because they are used in limited positions. For a child who uses voiced stops only in initial position and voiceless stops only in final position, for example, “bad,” “bat,” “pad,” and “pat” are homonyms because they are all pronounced as [bet]. Conversely, some children with disordered phonology may maintain a contrast, but without the expected feature. A child who does not voice final stops may lengthen the preceding vowel to indicate voicing: “bat” as [bæt] and “bad” as [bæd]. Such a child has phonological “knowledge” of the contrast and may independently develop voicing skills. Therefore, Elbert and Gierut (1986) suggest that features of which children have little knowledge are a higher priority for intervention. In other studies (e.g., Rvachew and Nowak, 2001), however, subjects have made more progress when most-knowledge features were addressed first.

Often, distinctive features are remediated with an emphasis on contrast, through minimal pair therapy. Treatment focuses on words differing by one distinctive feature. For example, the continuant feature could be taught by contrasting “tap” versus “sap” and “met” versus “mess.” Typically the sound that the child substitutes for the target sound (e.g., [t] for /s/) is compared to the target sound ([s]). Therapy may begin with discrimination activities: the child indicates the picture that corresponds to the word produced by the clinician (e.g., a hammer for “tap” versus an oozing tree for “sap”). This highlights the confusion that may result if the
wrong sound is used. Communication-oriented production activities are designed to encourage the child to produce the feature that had been missing from his system (e.g., to say “mess” rather than “met”). (See PHONOLOGICAL AWARENESS INTERVENTION FOR CHILDREN WITH EXPRESSIVE PHONOLOGICAL IMPAIRMENTS for a discussion of approaches in which the contrastive role of phonological features and structures is even more explicitly addressed.)

Gierut (1990) has tested the use of maximal pair therapy, in which the contrasting sounds differ on many features (e.g., [s] versus [z]). She has found that children may be able to focus better on the missing feature (e.g., continuant) when it is not contrasted with the substituting feature (e.g., noncontinuant) than when it is.

Markedness is another phonological concept with implications for remediation of sounds. Markedness reflects ease of production and perception: [θ] is marked, due to low perceptual salience; [b] is unmarked because it is easy to produce and to perceive. Children typically acquire the least marked sound classes (e.g., stops) and structures (e.g., open CV syllables) of their languages first (Dinnsen et al., 1990). It is unusual for a child to master a more marked sound or structure before a less marked one. Gierut’s (1998) research suggests that targeting a more marked structure or sound in therapy facilitates the acquisition of the less marked one “for free.” However, other researchers (e.g., Rvachew and Nowak, 2001) report more success addressing less marked sounds first.

Therapy based on nonlinear phonology stresses the importance of syllable and word structures as well as segments (Bernhardt, 1994). Minimal pair therapy is often used to highlight the importance of structure (e.g., “go” versus “goat,” “Kate” versus “skate,” “monkey” versus “monk”). The goal of this therapy is to expand the child’s phonotactic repertoire. For example, a child who previously omitted final consonants may or may not produce the correct final consonant in a given word, but will produce some final consonant. Structural and segmental deficits often interact in such a way that a child can produce a sound in certain positions but not others (Edwards, 1996).

The approaches described above focus primarily on what’s missing from the child’s phonological system. Another set of approaches focuses on what’s happening instead of the target production. For example, a child whose phonological system lacks the /θ, δ, s, z/ fricative phonemes may substitute stops (stopping), substitute lateral fricatives ([f, v]; fronting), substitute palatal fricatives ([ʃ, ʒ]; backing or palatalization), or omit /θ, δ, s, z/ in various word positions (initial consonant deletion, final consonant deletion, consonant cluster reduction). These patterns are referred to as phonological processes in speech-language pathology.

Phonological process therapy addresses three types of error patterns:

- Substitution processes: sounds with a certain feature are substituted by sounds with a different feature (e.g., fricatives produced as stops, in stopping; liquids produced as glides, in gliding).
- Phonotactic processes: sounds or syllables are omitted, added, or moved. The process changes the shape of the word or syllable. As examples, a CVC word becomes CV; a CCVC word becomes CVC, CVCVC, or CVCC; “smoke” becomes [moks].
- Assimilation processes: two sounds or two syllables become more alike. For example, a child who does not typically front velar consonants in words such as “go” may nonetheless say [dɔd] for “dog.” The final velar consonant becomes alveolar in accord with the initial consonant. Similarly, a two-syllable word such as “popcorn” may be reduplicated as [kɔkɔ]; the first syllable changes to match the second.

As in distinctive feature therapy, in phonological process therapy classes of sounds or structures that pattern together are targeted together. Again, either the entire class can be directly addressed in therapy or some representative members of the class may be selected for treatment, in the expectation that treatment effects will generalize to the entire class. The goal is to reduce the child’s use of that process, with resultant changes in her phonological system. For example, if the child’s phonological system is expanded to include a few final consonants, it is expected that she or he will begin to produce a variety of final consonants, not just those that were targeted in therapy.

Some therapists use traditional production activities (beginning at the word level) to decrease a child’s use of a phonological process; others use a minimal pair approach, comparing the targeted class (e.g., velars) with the substituting class (e.g., alveolars). The cycles approach (Hodson and Paden, 1991) to process therapy has some unique features. First, each session includes a period of auditory bombardment, during which the child listens (passively) to a list of words that contain the targeted sound class or structure. Second, each pattern is the focus for a predetermined length of time, regardless of progress. Then, treatment moves on to another pattern. Hodson and Paden argue that cycling is more similar to the phonological development of children without phonological disorders.

In summary, the goal of linguistically based approaches to phonological therapy is to make as broad an impact as possible on the child’s phonological system, making strategic choices of treatment goals that will trigger changes in untreated as well as treated sounds or structures.

—Shelley L. Velleman

References


Speech Disorders in Children: Motor Speech Disorders of Known Origin

By definition, children with a communication diagnosis of motor speech disorder have brain dysgenesis or have sustained pre-, peri-, or postnatal damage or disease to the central or peripheral nervous system or to muscle tissue that impairs control of speech production processes and subsequent actions of the muscle groups used to speak (respiratory, laryngeal, velopharyngeal, jaw, lip, and tongue) (Hodge and Wellman, 1999). This impairment may manifest with one or more of the following: weakness, tone alterations (hypertonia, hypotonia), reduced endurance and coordination, and involuntary movements of affected speech muscle groups (dysarthria), or as difficulty in positioning muscle groups and sequencing their actions to produce speech that cannot be explained by muscle weakness and tone abnormalities (apraxia of speech). Disturbances affecting higher mental processes of speech motor planning and programming underlie the motor speech diagnosis, apraxia of speech. To date, apraxia of speech of known origin in childhood is rare. Conditions in which it may appear are seizure disorders (e.g., Landau-Kleffner syndrome; Love and Webb, 2001), focal ischemic events (Murdoch, Ozanne,
and Cross, 1990), and traumatic brain injury. In these cases the speech apraxia is typically accompanied by expressive and receptive language deficits. Oropharyngeal apraxia and mutism have been reported following posterior fossa tumor resection in children (Dailey and McKhann, 1995). Depending on the severity and duration of the neurological insult, signs of childhood-acquired apraxia of speech remit and may disappear.

Disturbances affecting the execution of speech actions are diagnosed as dysarthrias, with subtypes identified by site of lesion, accompanying pathophysiological signs, and effects on speech production. Dysarthrias are the more common type of childhood motor speech disorder. A known neurological condition affecting neuromuscular function, including that of muscles used in speech, is a key factor leading to the diagnosis of dysarthria. A useful taxonomy of subtypes of childhood dysarthrias (with associated sites of lesion) was described by Love (2000) and includes spastic (upper motor neuron), dyskinetic (basal ganglia control circuit), ataxic (cerebellar control circuit), flaccid (lower motor neuron and associated muscle fibers), and mixed (two or more sites in the previous categories); the mixed type is the most common.

Limited research has been published on the nature of neuromuscular impairment in the various childhood dysarthrias and how this correlates with perceived speech abnormalities (Workinger and Kent, 1991). Solomon and Charron (1998) reviewed the literature on speech breathing in children with cerebral palsy. Love (2000) summarized the literature on respiratory, laryngeal, velopharyngeal, tongue, jaw, and lip impairment in children with dysarthria. It is difficult to generalize from the literature to individual cases because of the relatively small numbers of children studied and the range of individual differences in children with neurogenic conditions, even when they share the same neurological diagnosis. Furthermore, it cannot be assumed that because a neurological diagnosis has implicated a certain site of lesion (e.g., cranial nerve or muscle group), other muscle groups are not also impaired. Murdoch, Johnson, and Cross (1990) described the case of a child with Möbius syndrome (commonly held to result from damage to cranial nerves VI and VII) and also identified impaired function at the level of the velopharyngeal, laryngeal, and respiratory subsystems using perceptual and instrumental evaluation. However, a reduction in maximal range of performance of speech muscle groups, persistent dependencies between muscle groups (e.g., lip with jaw, jaw with tongue, tongue body with tongue tip), and reductions in speed, precision, and consistency of speech movements are common themes in the literature on the nature of impairment in childhood motor speech disorders.

The effects of most childhood motor speech disorders are to reduce the rate and quality of affected children’s speech development, frequency of speech use, speech intelligibility, speaking rate, and overall speech acceptability. The speech disorder can range from mild to so severe that the child never gains sufficient control over speech muscles to produce voice or recognizable speech sounds. These children’s psychosocial development is also at risk because of limitations imposed by the speech disorder on social interactions, which in turn may limit their academic progress because of fewer opportunities to gain experience using language (Hodge and Wellman, 1999).

When the cause of a childhood motor speech disorder is known, it is typically because a physician specializing in pediatric neurology has diagnosed it. If this neurological diagnosis (e.g., cerebral palsy, Möbius syndrome, muscular dystrophy) is made during the prelinguistic period, there is some expectation that if neuromuscular dysfunction was observed in earlier nonspeech activities (e.g., sucking, chewing, swallowing, control of saliva, facial expressions) of muscle groups that are involved in speech production, speech development will also be delayed and disrupted (Love, 2000). Congenital supra-bulbar paresis, or Worster-Drought syndrome, which has been classified by Clark et al. (2000) as a mild spastic quadriplegic type of cerebral palsy resulting from damage to the perisylvian areas of the cortex, often is not diagnosed until the child is older, if at all. Early diagnosis of this condition is important to speech-language pathologists (Crary, 1993; Clark et al., 2000) because it is predominantly characterized by persisting signs of abnormal neuromuscular dysfunction in oropharyngeal muscles during infants’ feeding and swallowing and later control of speech movements. Closer examination of children with congenital supra-bulbar paresis also reveals evidence of persisting neuromuscular abnormalities affecting gross and fine motor development and learning difficulties. These children need coordinated, multidisciplinary services like those afforded children with other subtypes of cerebral palsy.

In other cases of childhood motor speech disorders, no abnormal neurological signs are observed in the child’s early development, and signs of the motor speech disorder may be the first or only indication of neurological abnormality (Arvedson and Simon, 1998). In some of these cases, subsequent neurological investigation with electromyography, electroencephalography, neural imaging procedures such as magnetic resonance imaging, and metabolic testing identifies a neurological condition or lesion as the cause of the speech disturbance (e.g., seizures and brain dysmorphology in the bilateral perisylvian region, infection, tumor, progressive conditions such as facioscapulohumeral muscular dystrophy) (Hodge and Wellman, 1999). In rare cases, a motor speech disorder may result from treatment for another medical condition, such as surgery for cerebellar tumors or drug treatment for a debilitating movement disorder such as Tourette’s syndrome. In still other cases neurological investigation reveals no identifiable cause.

Many conditions that result in childhood motor speech disorders (e.g., cerebral palsy, traumatic brain injury, chromosomal abnormalities) also affect other areas of brain function. Therefore, children may show a mixed dysarthria or characteristics of both dysarthria and apraxia of speech, and they have a high probability...
of comorbidities affecting higher cognitive functions of language, thought, attention, and memory, sensory-perceptual processes and control of other motor systems (e.g., eyes, limbs, trunk, head).

Multidisciplinary assessment by members of a pediatric rehabilitation team is accepted clinical practice to determine the presence and severity of comorbid conditions that may negatively affect the child’s development. Specific to the motor speech disorder, assessment goals may include one or more of the following: establishing a differential diagnosis; identifying the nature and severity of impaired movement control in each of the muscle groups used to produce speech; describing the nature and extent of limitations imposed by the impairment on the child’s speech function in terms of articulatory adequacy, prosody and voice, and speech intelligibility, quality, and rate; determining the child’s ability to use speech, together with other modes, to communicate with others in various contexts of daily life; and making decisions about appropriate short- and long-term management (Hodge and Wellman, 1999; Yorkston et al., 1999; Love, 2000). At young ages or in severe cases, making a differential diagnosis can be difficult if the child has insufficient speech behaviors to analyze and lacks the attentional, memory, and cognitive abilities to execute tasks that are classically used to differentiate dysarthria from praxis disturbances. Hodge (1991) summarized strategies for assessing speech motor function in children 0–3 years old. Love and Webb (2001) reviewed primitive and oropharyngeal reflexes that may signal abnormal motor development at young ages. Hayden and Square (1999) developed a standardized, normed protocol to aid in the systematic assessment of neuromotor integrity of the motor speech system at rest and when engaged in vegetative and volitional nonspeech and speech tasks for children ages 3–12 years. Their protocol was designed to identify and differentially diagnose childhood motor speech disorders and is built on their seven-stage hierarchical model of speech motor development and control. Thoenen et al. (1999) described the use of several maximal performance tasks (diadochokinetic rates for repeated mono- and trisyllables, sustained vowel and fricatives) in a decision tree model to identify and diagnose motor speech disorders; they normed their model on children age 6 years and older.

Assessment of impairment should include evaluation of the integrity of structural as well as functional aspects of the speech mechanism, because abnormal resting postures and actions of oropharyngeal muscles can lead to abnormalities in the dental arches, poor control of oral secretions, and an increased risk for middle ear infections. The use of detailed, comprehensive protocols to guide assessment of speech mechanism impairment and interpretation of findings is recommended (e.g., Dworkin and Culatta, 1995; St. Louis and Ruscello, 2000). Instrumental procedures that assess the function of the various speech subsystems (respiratory, laryngeal, velopharyngeal, and oral articulatory) also help to determine the nature and severity of impairment if child cooperation allows (Murdoch, Johnson, and Theodoros, 1997). Procedures for assessing various aspects of speech function (e.g., phonological system and phonetic adequacy, intelligibility, prosody) and speech ability that can be repeated across time to index change are described in Hodge and Wellman (1999) and Yorkston et al. (1999). In addition to perceptual measures, acoustic measures such as second formant onset and extent and vowel area have been shown to be sensitive to both effective and ineffective compensatory articulation strategies used by children with motor speech disorders (e.g., Nelson and Hodge, 2000).

When making decisions about intervention, the clinician also needs to assess the child’s feeding behavior, cognitive and receptive and expressive language skills, hearing, psychosocial status and motivation to speak and communicate, gross and fine motor skills, general health and stamina, and the child’s family situation, including family members’ goals and perspectives. A multidisciplinary team approach to assessment and intervention planning, with ongoing involvement of the family in selecting, coordinating, and evaluating treatment service options, is critical to the successful management of children with motor speech disorders (Mitchell and Mahoney, 1995).

At present, most childhood motor speech disorders are considered to be chronic because they are the result of brain damage or dysgenesis, for which there is no cure. It is expected that the neurological diagnosis associated with the motor speech disorder will be chronic and may affect the child’s academic, social, and vocational future. However, appropriate and sufficient treatment can significantly improve these children’s communication effectiveness (Hodge and Wellman, 1999; Love, 2000). The overall goal of treatment is to help these children communicate in the most successful and independent manner possible and includes helping them become desirable communication partners. It is unrealistic to expect that these children will be provided with or benefit from continuous treatment from infancy to adulthood. Families have identified the preschool and early school years, and school transitions (i.e., change in schools due to family relocation, elementary to junior high, junior to senior high, and then to college), as times when the knowledge and support of speech-language pathologists is particularly needed. Children with a congenital or early-onset condition that results in a motor speech disorder must learn how to produce the phonological system of their language with reduced control of the muscle groups used to produce and shape the sound contrasts that signal meaning to their listeners. Explicit, goal-directed opportunities for extensive practice in producing and combining speech sounds in meaningful utterances are typically required for these children to attain their potential for learning their phonological system, its phonetic realization, and making their speech intelligible. In addition to speech development, the child’s language, cognitive, and social development is also at risk because of the important role that speech plays in development. Children who suffer neurological
insight after the primary period of speech development has occurred (e.g., after 3–4 years of age) are faced with the task of relearning a system that has been acquired with normal speech motor control, so they have internal models of their phonological-phonetic system in place. For these children, the focus is on relearning these, and then monitoring and providing support as needed for acquisition of new spoken language skills that had not been acquired at the time of the neurological insult.

Treatment planning for all children with motor speech disorders should include a team approach that promotes active family involvement in making decisions and implementing treatment, attention to principles of motor learning at a level that is developmentally appropriate to the child (Hodge and Wellman, 1999), consideration of a variety of service delivery forms, and a holistic view of the child with a motor speech problem (Mitchell and Mahoney, 1995). Training tasks need to be goal-directed and should actively engage and involve the child as a problem solver. Because learning is context-specific, training activities should simulate real-world tasks as much as possible and be enjoyed by the child. Training goals should build on previously learned skills and behaviors. The child must have multiple opportunities to practice attaining each goal, and should have knowledge of results.

A combination of treatment approaches that address multiple levels of the communication disorder (impairment, speech functional impairment, and ability to communicate in various contexts) is typically used in management programs for children with motor speech disorders. The particular approaches change with the child’s and family’s needs and as the child’s abilities change. Possible approaches include the following (Hodge and Wellman, 1999; Yorkston et al., 1999; Love, 2000):

**General**
- Educate family members, other caregivers, and peers about the nature of the child’s speech disorder and ways to communicate effectively with the child.
- Augment speech with developmentally appropriate alternative communication modes.
- Provide receptive and expressive language treatment (both spoken and written) as appropriate, and integrate this with speech training activities when possible.
- Address related issues as necessary, including management of any interfering behaviors (e.g., attention, lack of motivation), control of oral secretions and swallowing, oral-dental status, and sensory (auditory-visual) status.

**Speech-Specific**
- Increase the child’s physiological support for speech by increasing spatiotemporal control and coordination of speech muscle groups (respiratory, laryngeal, velopharyngeal, lips, tongue, jaw). The objective is to increase movement control for speech production, and the selection, implementation, and evaluation of these techniques must be made relative to their effect on increasing speech intelligibility and quality. This includes working with other members of pediatric rehabilitation teams to optimize the child’s seating and positioning for speech production (Solomon and Charron, 1998; Love, 2000).
- Develop the child’s phonological and phonetic repertoire with attention to level of development as well as the specific profile of muscle group impairment; also, develop the child’s phonological awareness and pre-literacy and literacy skills.
- Improve vocal loudness and quality through behavioral training.
- Increase speech intelligibility and naturalness through prosodic training.
- Use behavioral training to maximize the effects of drug treatments, prosthetic compensation (e.g., palatal lift), or surgery (e.g., pharyngoplasty).
- Identify and promote the use of effective speech production compensatory behaviors.

**Communication Effectiveness**
- Teach the effective use of interaction enhancement strategies.
- Model and promote the use of effective conversational repair strategies and speech production self-monitoring skills.
- Teach effective cognitive strategies so the child can use word choice and syntactic structure to maximize listeners’ comprehension.
- Promote maintenance of speech production skills that have been established and self-monitoring of communication skills.
- Implement strategies to increase the child’s self-esteem and self-confidence in initiating and participating in communication interactions.

See also DEVELOPMENTAL APRAXIA OF SPEECH; DYSPHASIC ARTHRIAS: CHARACTERISTICS AND CLASSIFICATION; MOTOR SPEECH INVOLVEMENT IN CHILDREN.

—Megan M. Hodge

**References**


**Further Readings**


**Speech Disorders in Children: Speech-Language Approaches**

Children with speech disorders often display difficulty in other domains of language, suggesting that they experience difficulty with the language learning process in general. A theoretical shift from viewing children’s speech disorders as articulatory-based to viewing them from a linguistic perspective was precipitated by the application of phonological theories and principles to the field of speech pathology, beginning in the late 1970s. Implicit in this shift was the recognition that acquisition of a linguistic system is a gradual, primarily auditory-perceptually based process involving the development of receptive knowledge first (Hodson and Paden, 1991; Ingram, 1997). Another implication from research on normal phonological acquisition is that linguistic input should demonstrate a sound’s contrastive role in the ambient phonology because what children are developing is phonological oppositions. As a result of applying a linguistic model to intervention, speech-language approaches have flourished in the past two decades for treating children’s speech disorders. These approaches are characterized by: (1) an emphasis on the function of the phonological system to support communication,
and thus on the pragmatic limitations of unintelligible speech; (2) a focus on the contrastive nature of phonemes and the use of minimal pair contrast training to facilitate reorganization of the system; (3) the use of phonological analysis and description to identify error patterns; (4) the selection of error patterns for elimination and of sound classes and features for acquisition; and (5) the use of a small set of sounds as exemplars of those patterns/features for acquisition. What separates speech-language approaches from other articulation approaches is the use of phonological analysis to identify error patterns affecting sound classes and sequences rather than the selection of isolated sounds to be trained in each word position. Speech-language approaches differ from articulation approaches in their focus on modifications of groups of sounds via a small set of exemplars (e.g., /f, s/ to represent all fricatives) and their emphasis on contrastivity for successful communication in a social context.

As such, minimal pair contrast treatment is a speech-language approach that highlights the semantic confusion caused when the child produces a sound error that results in a pair of homonyms (e.g., “sun” and “ton” both produced as /tan/). This technique involves contrasting a pair of words in which one word contains the child’s error production and the other contains the target production (with phonemes differing by only one feature). In minimal pair approaches, the child is instructed to make perceptual and productive contrasts involving the target sound and his or her error. The goal of treatment is to help the child learn to produce the target sound in the word pair to signal a difference in meaning between the two words. Minimal pair contrast interventions have been shown to be effective in eliminating error patterns and increasing the accuracy of target and related error sounds (Ferrier and Davis, 1973; Elbert, Rockman, and Saltzman, 1980; Blache, Parsons, and Humphreys, 1981; Weiner, 1981; Elbert, 1983; Tyler, Edwards, and Saxman, 1987; Saben and Ingham, 1991).

Minimal pair approaches that involve solely perception of word contrasts, however, have not resulted in as much change as when production practice with models and phonetic cues is also included (Saben and Ingham, 1991).

Minimal pair contrasts are used somewhat differently in Metaphon (Howell and Dean, 1994; Dean et al., 1995), a “cognitive-linguistic” approach. This approach is considered cognitive-linguistic because it facilitates conceptual development and cognitive reorganization of linguistic information. Its aim is to increase, at the metalinguistic level, awareness and understanding of sound class differences primarily through classification techniques, with little emphasis on production. For example, to contrast alveolars and velars, the concepts of front and back are introduced and applied first to nonspeech sorting and classification activities. Next, these concepts are transferred to the speech domain by having the child listen to minimal pairs and judge whether or not words begin with front or back sounds. Dean et al. (1995) present evidence from several children suggesting that this approach effectively reduced the application of selected phonological processes.

The cycles approach, proposed by Hodson and Paden’s (1991), involves a goal attack strategy that capitalizes on observations concerning the gradual manner in which normally developing children acquire their phonological systems. Groups of sounds affected by an error pattern are introduced for only 1 or 2 weeks, then a new error pattern is introduced. Thus, the criterion for advancement to a new goal or target is time-based rather than accuracy-based. A cycle can range from 5 to 15 weeks, depending on the number of deficient patterns, and once completed, the sequence is recycled for the error patterns that still remain in the child’s speech. Hodson and Paden’s cycle approach involves auditory bombardment and production practice in the form of picture- and object-naming activities, and reportedly eliminates most of a child’s phonological error patterns in 1–2 years of intervention (Tyler, Edwards, and Saxman, 1987; Hodson and Paden, 1991; Hodson, 1997).

In contrast to the three approaches just described, which focus on speech within a linguistic framework, there are language-based approaches in which little attention may be drawn to sound errors and these errors may not be specific targets of intervention. Instead, the entire language system (syntax, semantic, phonology, pragmatics) is targeted as a tool for communication, and improvements in phonology are expected from a process of “whole to part learning” (Norris and Hoffman, 1990; Hoffman, 1992). Phonological changes might be expected to occur because phonemes are practiced as parts of larger wholes within the script for an entire event. Language-based approaches involve a variety of naturalistic, conversationally embedded techniques such as scaffolding narratives, focused stimulation in the form of expansions and recasts, and elicited production devices such as forced-choice questions, cloze tasks, and preparatory sets.

Norris and Hoffman’s (1990, 1993) language-based approach focuses on scaffolding narratives in the form of expansions, expatiations, and turn assistance devices to help the child talk about picture sequences with higher levels of discourse and semantic complexity. Hoffman, Norris, and Monjure (1990) contrasted their scaffolded narrative approach to a phonological process approach in two brothers with comparable phonological and language deficits. The narrative intervention facilitated gains in phonology that were similar to the phonological approach, and greater gains in syntactic, semantic, and pragmatic performance.

Other language-based approaches reported in the literature have focused primarily on morphosyntactic goals (e.g., finite morphemes, pronouns, complex sentences) using focused stimulation designed to provide multiple models of target morphosyntactic structures in a natural communicative context. Procedures have involved recasts and expansions of child utterances, and opportunities to use target forms in response to forced-choice questions, sentence fill-ins, requests for elaboration, or false assertions in pragmatically appropriate contexts.
(Cleave and Fey, 1997). Researchers have been interested in the cross-domain effects of these procedures on improvement in children’s speech disorders. Fey et al. (1994) examined the effects of language intervention in 25 children with moderate to severe language and speech impairments who were randomly assigned to a clinician treatment group, a parent treatment group, or a delayed-treatment control group. The treatment groups made large gains in grammar after 5 months of intervention, but improvement in speech was no greater than that achieved by the control group. Tyler and Sandoval (1994) examined the effects of treatment focused only on speech, only on morphosyntax, and on both domains in six preschool children. The two children who received morphosyntactic intervention showed improvements in language but negligible improvement in phonology.

In contrast to these findings that language-based intervention focused on morphosyntax does not lead to gains in speech, Tyler, Lewis, Haskell, and Tolbert (2002) found that a morphosyntax intervention addressing finite morphemes led to improvement in speech in comparison to a control group. Tyler et al. (2002) investigated the efficacy and cross-domain effects of both a morphosyntax and a phonological intervention. Ten preschool children were assigned at random to an intervention of two 12-week blocks, beginning with either a block focused on speech first or a block focused on morphosyntax first. Treatment efficacy was evaluated after one block in the sequence was applied. Not only was the morphosyntax intervention effective in promoting change in morphemes marking tense and agreement in comparison to the no-treatment control group, but it led to improvement in speech that was similar to that achieved by the phonology intervention. Thus, for children who received language intervention, the amount of speech improvement was significantly greater than that observed for the control group. In a similar study, Matheny and Panagos (1978) examined the effect of highly structured interventions focused on syntax only and articulation only in children with deficits in both domains, compared with a control group. Each group made significant gains in the treated domain in comparison with the control group, but also made improvements in the untreated domain. Thus, a language-based intervention focused on complex sentence structures led to improved speech.

Findings regarding the effects of language intervention on speech are equivocal, particularly results from methodologically rigorous studies with control groups (Matheny and Panagos, 1978; Fey et al., 1994; Tyler et al., 2002). One variable that may account for these differing results is the use of different measures to document change. Matheny and Panagos used before and after standardized test scores, whereas Fey et al. used Percent of Consonants Correct (PCC; Shriberg and Kwiatkowski, 1982), a general measure of consonant accuracy, and Tyler et al. used a more discrete measure of target and generalization phoneme accuracy. Nonetheless, the collective findings from studies of the effects of different language-based approaches on speech suggest that some children, especially those with both speech and language impairments, will show improvement in speech when the intervention focuses on language. Determining exactly who these children are is difficult. Preliminary evidence suggests that children whose phonological systems are highly inconsistent may be good candidates for a language-based approach (Tyler, 2002). Finally, service delivery restrictions may dictate the use of language-based approaches in classroom or collaborative settings. These approaches deserve further investigation for their possible benefit in remediating both speech and language difficulties.

In summary, a variety of speech-language approaches have been shown to be effective in improving speech intelligibility and reducing the number and severity of error patterns in children with speech disorders. Although these approaches employ different teaching methods, they originate in a linguistic model and share an emphasis on the function of the phonological system to support communication, and on the contrastive nature of phonemes to reduce the pragmatic limitations of unintelligible speech.

—Ann A. Tyler

References


**Further Readings**


**Speech Disorders Secondary to Hearing Impairment Acquired in Adulthood**

Hearing loss is very common in the general population, with a prevalence of 82.9 per 1000 (U.S. Public Health Service, 1990). It becomes more common with age as a result of noise exposure, vascular disease, ototoxic agents, and other otological diseases. After arthritis and hypertension, hearing loss is the third most common...
chronic condition in persons over 65 (National Center for Health Statistics, 1982). In a study of 3556 adults from Beaver Dam, Wisconsin, Cruickshanks et al. (1998) found prevalence rates for hearing loss of 21% in adults ages 48–59 years, 44% for those ages 60–69, 66% for those ages 70–79, and 90% for those ages 80–92. The prevalence of perceived hearing handicap, however, is lower than the true prevalence of hearing loss. By age 70, approximately 30% of the population perceives themselves as hearing impaired, and by 80 years, 50% report being hearing impaired (Desai et al., 2001). There is also some indication that the prevalence of hearing impairment in persons 45–69 years old is increasing, especially among men (Wallhagen et al., 1997).

The typical hearing loss configuration in adults is a bilateral high-frequency sensory loss with normal or near normal hearing in the low frequencies (Moscicki et al., 1985; Cruickshanks et al., 1998). Men tend to have more hearing loss than women, and white individuals report greater hearing impairment than African Americans (Cruickshanks et al., 1998; Desai et al., 2001).

Although hearing loss is common in the general population, its effects on speech production are most pronounced in individuals who have congenital hearing loss or hearing losses acquired in early childhood. For individuals who acquire hearing loss as adults, the impact on speech production is limited and usually does not result in any perceptible speech differences (Goehl and Kaufman, 1984). The preservation of speech in most adults with hearing loss likely is a consequence of residual hearing sufficient for auditory feedback.

Speech differences have, however, been reported for some persons with complete loss or nearly complete loss of hearing. These individuals tend to remain intelligible, although the speaking rate may be reduced by about a third when compared with normal-hearing speakers (Leder, Spitzer, Kirchner, et al., 1987). A decreased speaking rate is reflected in increased sentence and pause durations as well as increased word, syllable, and vowel durations (Kirk and Edgerton, 1983; Leder et al., 1986; Leder, Spitzer, Kirchner, et al., 1987; Waldstein, 1990; Lane et al., 1998). Movement durations associated with articulatory gestures also are prolonged in some adventitiously deafened adults (Matthies et al., 1996), and it has been suggested that this overall decrease in rate contributes to a reduction in speech quality and communication effectiveness (Leder, Spitzer, Kirchner, et al., 1987).

Changes in respiratory and vocal control have been noted, as evidenced by abnormal airflow, glottal aperture, and air expenditure per syllable as well as frequent encroachment on respiratory reserve (Lane et al., 1991, 1998). Adventitiously deafened adults also tend to exhibit increased breathiness, vocal intensity, and mean fundamental frequency (Leder, Spitzer, and Kirchner, 1987; Leder, Spitzer, Milner, et al., 1987; Lane et al., 1991; Lane and Webster, 1991; Perkell et al., 1992). In addition, the fundamental frequency tends to be more variable, particularly on stressed vowels (Lane and Webster, 1991).

Reduced phonemic contrast also characterizes the speech of some adventitiously deafened adults. Lane et al. (1995) observed that voice-onset time tends to decrease for both voiced and voiceless stop consonants, while Waldstein (1990) observed this effect only with voiceless stop consonants. Vowel, plosive, and sibilant spectra become less distinct, and vowel formant spacing for some speakers becomes more restricted and centralized (Waldstein, 1990; Lane and Webster, 1991; Matthies et al., 1994, 1996; Lane et al., 1995). The first vowel formant commonly is elevated, with some speakers also exhibiting a reduction in second formant frequency (Perkell et al., 1992; Kishon-Rabin et al., 1999). A greater overlap in articulator postures and placements has also been observed, with a tendency for the consonant place of articulation to be displaced forward and vowel postures to be neutralized (Matthies et al., 1996). Fricatives and affricates appear particularly prone to deterioration with profound hearing loss (Lane and Webster, 1991; Matthies et al., 1996). Many of these changes, although subtle in many cases and variable in expression across this population, are consistent with the speech differences common to speakers with prelingual hearing loss. Although the evidence is limited, owing to the small numbers of subjects examined, the data across studies suggest that the effects of hearing loss on speech production are most pronounced if the hearing loss occurs in the teens and early twenties than if it occurs in later adulthood.

The primary management procedure for adults with acquired hearing loss severe enough to compromise speech is to restore some degree of auditory feedback. The initial intervention typically consists of fitting traditional amplification in the form of hearing aids. For first-time hearing aid wearers, postfitting rehabilitation consisting of counseling and auditory training improves auditory performance and retention of the hearing aid, although secondary benefit in respect to speech production in adults has not been studied systematically (Walden et al., 1981).

A sensory implant often is recommended for adults who do not receive sufficient benefit from hearing aids or who cannot wear hearing aids (see COCHLEAR IMPLANTS). Individuals with an intact auditory nerve and a patent cochlea are usually candidates for a cochlear implant. Persons who cannot be fitted with a cochlear implant, such as persons with severed auditory nerves or ossified cochleae, may be candidates for a device that stimulates the auditory system intracranially, such as a brainstem implant. As with hearing aids, adult patients receiving sensory implants benefit from pre- and postfitting counseling and frequent monitoring. With current technologies, many patients show substantial improvement in auditory and speech function within months of device activation with little or no additional intervention, although normalization of all speech parameters may never occur or may take years to achieve (Kishon-Rabin et al., 1999). Some speech parameters, such as vocal intensity and fundamental frequency, show some degree of reversal when an implant is temporarily turned off and...
then on, but the extent and time course of overall recovery after initial activation vary with the individual. The variability in speech recovery after initial implant activation appears to result from a number of factors, among them age at onset of hearing loss, improvement in auditory skills after implant activation, extent of speech deterioration prior to activation, and the speech parameters affected (Perkell et al., 1992; Lane et al., 1995, 1998; Kishon-Rabin et al., 1999; Vick et al., 2001). As a result, some patients may benefit from behavioral intervention to facilitate recovery. In particular, persons with poor speech quality prior to receiving an implant, as well as persons with central auditory deficits and compromised devices, might benefit from systematic auditory, speech, and communication skills training, although the relationship between speech recovery and behavioral treatment has received little investigation in these patients.

See also AUDITORY TRAINING; COCHLEAR IMPLANTS; HEARING LOSS SCREENING; THE SCHOOL-AGE CHILD; NOISE-INDUCED HEARING LOSS; OTOTOXIC MEDICATIONS; PRESBYCUSIS; SPECREADING TRAINING AND VISUAL TRACKING.

—Sheila Pratt

References


Further Readings

Speech Issues in Children from Latino Backgrounds

Research into English phonological development in typically developing children and children with phonological disorders has been occurring since the 1930s (e.g., Wellman et al., 1931; Hawk, 1936). There is limited information on phonological development in Latino children, particularly those who are monolingual Spanish speakers and bilingual (Spanish and English) speakers. Over the past 15 years, however, phonological information collected on monolingual Spanish speakers and bilingual (Spanish-English) speakers has increased greatly. This entry summarizes information on phonological development and disorders in Latino children, focusing on those who are Spanish-speaking. Spanish phonology and phonological development in typically developing Spanish-speaking children, Spanish-speaking children with phonological disorders, typically developing bilingual (Spanish-English) children, and bilingual (Spanish-English) children with phonological disorders will be reviewed.

There are five primary vowels in Spanish, the two front vowels, /i/ and /e/, and the three back vowels, /u/, /o/, and /a/. There are 18 phonemes in General Spanish (Núñez-Cedeño and Morales-Front, 1999): the voiceless unaspirated stops, /p/, /t/, and /k/; the voiced stops, /b/, /d/, and /g/; the voiceless fricatives, /f/, /x/, and /s/; the affricate, /ts/; the glides, /w/ and /j/; the lateral, /l/; the flap /r/, the trill /r/; and the nasals, /m/, /n/, and /n/. The three voiceless stops /b, d, g/ are in complementary distribution with the spirants /b̪/ (voiced bilabial), /d̪/ (voiced interdental), and /x̪/ (voiced velar), respectively. The spirant allophones most generally occur intervocally both within and across word boundaries (e.g., /dedo/ (finger) → /dedo/; /la boca/ (the mouth) → /la boka/) and in word internal consonant clusters (e.g., /ablar/ (to talk) → /aláfar/).

The phonetic inventory of Spanish differs from that of English. Spanish contains some sounds that are not part of the English phonetic system, including the voiced palatal nasal [n], as in [nipo] (boy), the voiceless bilabial fricative [ϕ], as in [emperomo] (sick), the voiceless velar fricative [x], as in [relaxo] (watch), the voiced spirants [b], as in [klaibo] (nail), and [x], as in [lavo] (lake), the alveolar trill [r], as in [pexo] (dog), and the voiced uvular trill [R], as in [Roto] (broken).

As in English, there are a number of dialectal varieties associated with Spanish. In the United States, the two most prevalent dialect groups of Spanish are Southwestern United States (e.g., Mexican Spanish) and Caribbean (e.g., Puerto Rican Spanish) (Iglesias and Goldstein, 1998). Unlike English, in which dialectal variations are generally defined by alterations in vowels, Spanish dialectal differences primarily affect consonants. Specifically, fricatives and liquids (in particular /s/, /t/, and /r/) tend to show more variation than stops, glides, or the affricate.

Common dialectal variations include deletion and/or aspiration of /s/ (e.g., /dos/ (two) → [do] or [doʰ]); deletion of /t/ (e.g., /kortar/ (cut) → [kotar]); substitution of /l/ or /j/ for /r/ (e.g., /kortar/ → /koltar/koiutar/); and substitution of /x/ or /R/ for /t/ (e.g., /pero/ (dog) → /pexo/prRo/). It should be noted that not every feature is always evidenced in the same manner and that not every speaker of a particular dialect uses each and every dialectal feature.

Phonological Development in Monolingual Spanish-Speaking Children. Most of the developmental phonological data on Spanish have been collected from typically developing, monolingual children. Data from segment-based studies suggest that typically developing,
were not usually observed in typically developing, exhibited by children with phonological disorders that frequency process (percentages of occurrence greater 15% and 30% for three processes) were initial consonant frequency processes (percentages of occurrence between 10% and 30% of possible occurrences). Goldstein and Iglesias (1993) found that low-frequency phonological processes by the time they reach 3½ years of age (e.g., A. Gonzalez, 1981; Mann et al., 1992; Goldstein and Iglesias, 1996a). Commonly occurring phonological processes (percentages of occurrence greater than 10%) included postvocalic singleton omission, stridency deletion, tap/trill /r/, consonant sequence reduction, and final consonant deletion. Less commonly occurring processes (percentages of occurrence greater than 10%) were fronting (both velar and palatal), prevocalic singleton omission, assimilation, and stopping.

Although there have been quite a number of studies characterizing phonological patterns in typically developing Spanish-speaking children, this information remains sparse for Spanish-speaking children with phonological disorders. Goldstein and Iglesias (1993) examined consonant production in Spanish-speaking preschoolers with phonological disorders and found that all stops, the fricative [f], the glides, and the nasals were produced accurately more than 75% of the time. The spirants [s] and [θ], the affricate, the flap [ɾ], the trill [r], and the lateral [l] were produced accurately 50%–74% of the time. Finally, the fricative [s], the spirant [s], and clusters were produced accurately less than 50% of the time.

Phonological development in Spanish-speaking preschool children with phonological disorders has also been examined (Meza, 1983; Goldstein and Iglesias, 1996b). Meza (1983) found that these children showed errors on liquids, stridents, and bilabials in more than 30% of possible occurrences. Goldstein and Iglesias (1996b) found that low-frequency phonological processes (percentages of occurrence less than 15%) were palatal fronting, final consonant deletion, assimilation, velar fronting, and weak syllable deletion. Moderate frequency processes (percentages of occurrence between 15% and 30% for three processes) were initial consonant deletion, liquid simplification, and stopping. The high-frequency process (percentages of occurrence greater than 30%) was cluster reduction. Other error types exhibited by children with phonological disorders that were not usually observed in typically developing, Spanish-speaking children were deaffrication, lisping, and backing.

Phonological Development in Bilingual (Spanish-English) Children. There is increasing evidence that the phonological systems of bilingual (English-Spanish) speakers develop somewhat differently from the phonological system of monolingual speakers of either language. Gildersleeve, Davis, and Stubbe (1996) and Gildersleeve-Neumann and Davis (1998) found that in English, typically developing, bilingual preschoolers showed an overall lower intelligibility rating, made more errors overall (on both consonants and vowels), distorted more sounds, and produced more uncommon error patterns than monolingual children of the same age. Gildersleeve, Davis, and Stubbe (1996) and Gildersleeve-Neumann and Davis (1998) also found higher percentages of occurrence (7%–10% higher) for typically developing, bilingual children (in comparison to their monolingual peers) on a number of phonological processes, including cluster reduction, final consonant deletion, and initial voicing. This discrepancy between monolingual and bilingual speakers, however, does not seem to be absolute across the range of phonological processes commonly exhibited in children of this age. Goldstein and Iglesias (1999) examined English and Spanish phonological skills in 4-, 5-, and 6-year-old, typically developing bilingual children and found that some phonological patterns (e.g., initial consonant deletion and deaﬀrication) were exhibited at somewhat lower rates in bilingual children than in monolingual children with phonological disorders than has been reported for monolingual, Spanish-speaking children with phonological disorders (Goldstein and Iglesias, 1996b). Thus, although the average percentage-of-occurrence diﬀerence is not large between monolingual and bilingual speakers, the results indicate that bilingual children will not always exhibit higher percentages of occurrence on phonological processes than monolingual children.

Goldstein and Washington (2001) indicated that the phonological skills of 4-year-old bilingual children were similar to their monolingual counterparts; however, the substitution patterns used for the target sounds flap /ɾ/ and trill /ɾ/ did vary somewhat between bilingual and monolingual speakers. For example, in bilingual children [ɾ] was a common substitute for the trill, but it was a relatively rare substitute for the trill in monolingual children. All four studies also found that bilingual children exhibited error patterns found in both languages (e.g., cluster reduction) as well as those, like liquid glide, that were typical in one language (English) but atypical in the other (Spanish).

Data from bilingual children with phonological disorders indicate that they exhibit more errors, lower rates of accuracy on consontants, and higher percentages of occurrence for phonological processes than either typically developing, bilingual children (Goldstein and Iglesias, 1999) or monolingual, Spanish-speaking children with phonological disorders (Goldstein and Iglesias, 1996b). The types of errors exhibited by the children, however, are similar regardless of target language (i.e., Spanish versus English). Specifically, bilingual children with phonological disorders showed higher error rates on clusters, fricatives, and liquids than other classes of sounds. Finally, percentages of occurrence for phono-
logical processes were higher overall for bilingual children with phonological disorders than for monolingual, Spanish-speaking children with phonological disorders.

The number of Latino children who speak Spanish continues to increase. Developmental phonological data collected from typically developing, monolingual Latino children who speak Spanish indicate that by age 3½, they use the dialectal features of their speech community and have mastered the majority of sounds in the language. The phonological development of typically developing, bilingual (Spanish-English) speakers is somewhat different from that of monolingual speakers of either language.

— Brian Goldstein

References


Further Readings


Speech Sampling, Articulation Tests, and Intelligibility in Children with Phonological Errors

Children who have speech sound disorders can reasonably be separated into two distinct groups. One group comprises children for whom intelligibility is a primary issue and who tend to use many phonological processes, especially deletion processes. These children are generally in the preschool age range. The second group includes children who have residual errors, that is, substitution and distortion errors that are relatively few in number. These children are typically of school age, and intelligibility is better than in the first group. The first group, namely children with phonological difficulties, is the focus of this entry.

In most cases, when we assess speech production in a preschool child, the purposes of the assessment are to describe the child’s phonological system and make decisions about management, if needed. A good audio- or video-recorded speech sample from play interactions with parents or the clinician can capture all of the primary data needed to describe the system, or it can be supplemented with a single-word test instrument. Typically, we define an adequate conversational speech sample as one that includes at least 100 different words (Crystal, 1982). Additionally, these words should not be direct or immediate repetitions of an adult model. Finally, for children with poor intelligibility, it is helpful for the examiner to repeat what he believes the child said after each utterance so that this spoken gloss is also recorded on the tape.

Once the recording has been made, the next task is for the examiner to gloss and transcribe the speech sample, ideally using narrow transcription (Edwards, 1986; Shriberg and Kent, 2003). Professionals who frequently do extensive transcription may wish to use a computer-based system. (See Masterson, Long, and Buder, 1998, for an excellent review of such software.) However, such systems are only as good as the clinician’s memory for the symbols and diacritics and her memory of their location on the keyboard; consequently, doing transcription by hand may be the more reliable way to go about this task if one does not do it frequently.

With the transcript in hand, the clinician now has a choice of types of analyses. First of all, one can undertake both independent and relational analyses (Stoel-Gammon and Dunn, 1985). Independent analyses treat the child’s system as self-contained, that is, with no reference to the adult system. They include a phonetic inventory for consonants and perhaps for vowels, as well as tallies of syllable or word shapes. Relational analyses, on the other hand, explicitly compare the child’s production to that of the adult, including a segmental (phonemic) inventory for consonants and perhaps for vowels and a list of the phonological processes that the child uses.

Independent analyses are appropriate for children who are very young, or who have very poor intelligibility, or who appear to use few differentiated speech sounds. Clinicians typically devise their own forms for these analyses, although some of the software mentioned above permits certain of these independent analyses to be done automatically. Frequency of use is an issue in independent analyses, so the various phones and syllable structures that appear in the transcript should be tallied on the inventory form. It is helpful to structure these forms into major consonant classes and major vowel classes, as well as syllable-initial, syllable-final, and intervocalic. In addition, for some types of analysis, separate inventories should be done for one-, two-, and three-syllable utterances or words.

One kind of relational analysis, the segmental or phonemic inventory, which compares the child’s production to the adult target, is more familiar to clinicians...
because it resembles typical published tests of articulation and phonology. Phonological process analysis is also considered to be relational in nature. If clinicians are working from a transcript of conversational speech, all of the software mentioned earlier can provide at least a list of phonological processes. Alternatively, clinicians may again devise their own forms for the segmental inventories and the list of phonological processes. Typically, the list of phonological processes will include the 8–10 processes commonly listed in texts and tests of phonology, as well as any unique or idiosyncratic processes that the child uses. The examiner then goes through the transcript noting what the child produces for each adult form. These productions are also tallied.

One other important measure that is relational in nature is the Percentage of Consonants Correct (PCC; Shriberg and Kwiatkowski, 1982). The number of adult targets that the child attempts is tallied, using the standards of colloquial speech, and the number of targets that the child produces acceptably is also tallied. Simple division of these tallies results in the PCC. The PCC is considered to be a measure of severity. It is a useful measure for assessing change over long periods of treatment, such as 6 months.

Tests of Phonology. Several tests published commercially permit analysis of children’s use of phonological processes on the basis of single-word naming of objects or pictures. They include the Bernthal-Bankson Test of Phonology (BBTOP; Bernthal and Bankson, 1990), the Khan-Lewis Phonological Analysis (KLPA; Khan and Lewis, 1986), and the Smit-Hand Articulation and Phonology Evaluation (SHAPE; Smit and Hand, 1997), all of which are based on pictures or photos, and the Assessment of Phonological Processes-Revised (APP-R; Hodson, 1986), which is based on object naming.

Tests of phonology can complement analyses of the conversational sample. One of their virtues is that for the phonological processes that are assessed, they incorporate multiple exemplars of each process, so that there is some assurance that the child’s use of the process is not happenstance. For some extremely unintelligible children, tests of phonology may be the only way to figure out the child’s patterns because the clinician at least knows what the intended word should be. Some of these tests (BBTOP, KLPA, and SHAPE) also permit comparisons to be made between the child’s performance and normative data.

On the other hand, if a test of phonology becomes the primary assessment tool, the clinician needs to be aware that measures derived from single-word naming may differ from measures derived from conversation. The conversational speech of disordered children generally includes more nonadult productions than does single-word naming. In addition, tests of phonology tend to deal with a very circumscribed set of phonological processes, and they do not deal at all with vowel productions. Consequently, if the child uses important but idiosyncratic processes, such as glottal replacement, or has systematic vowel errors, they may not be picked up. (However, SHAPE has an extensive list of potential idiosyncratic processes in an appendix, along with instructions for determining the frequency of use of these processes.)

Specialized testing using single-word stimuli may be required for specific treatment orientations. For example, in order to carry out a generative analysis, Elbert and Gierut (1986) have developed a test with more than 300 items in which bound morphemes, such as re-(meaning again) or the diminutive -y or -ie, are added to the word at either end. The purpose is to determine whether the child changes an error production of the first consonant or the last consonant in the presence of the morphological addition and changes it in a way that clarifies the child’s underlying representation.

Another example of specialized testing using either single words or connected speech is the elicitation of data needed for nonlinear analysis (Bernhardt and Stoe-Gammon, 1994). Nonlinear approaches deal with the hierarchies of representation of words, for example, at the segmental level, the syllable level, the foot level, and so on. Productions of multisyllabic words with varying stress patterns are needed to complete these analyses. In most cases these multisyllabic words must be elicited by imitation or picture naming.

Finally, an altogether different type of specialized testing is the determination of stimulability. A child is said to be stimulable for an error sound if the clinician can elicit an acceptable production using models, cues, or phonetic placement instructions. This part of the assessment is often performed informally and usually for just a few error sounds. However, Perrine, Bain, and Weston (2000) have devised a systematic way to assess stimulability based on a hierarchy of cues and models that is helpful in planning intervention.

Intelligibility. Intelligibility refers to how well the child’s words can be understood by others. There are at least two measures of intelligibility that are based on counts of intelligible words, as well as many scales for making perceptual judgments of intelligibility. The most straightforward numerical measure, Percent of Intelligible Words (PIW), has been described by Shriberg and Kwiatkowski (1982). To determine this measure, a person who does not know the child listens to the conversational speech sample but does not hear the clinician’s comments. This person attempts to gloss the sample. The number of words correctly glossed by the listener is divided by the total number of words to obtain the PIW.

A second measure based on counts, the Preschool Intelligibility Measure, has been developed by Morris, Wilcox, and Schooling (1995). The child is asked to imitate a series of one- or two-syllable words that are selected randomly from a large database of words, and her productions are recorded. Then the audiotape is played for listeners, who see 12 foils for each word and circle the one they think the child said. This measure is better suited for documenting changes in intelligibility over time than it is for initial evaluation.
Other scales of intelligibility involve judgments on the part of the clinician or significant others in the child’s environment about how well the child communicates. For example, teachers might be asked to rate how well they understand the child on a 6-point scale, with 1 representing “all the time” and 6 representing “never.” Or a parent might be asked to rate the difficulty that family members have in understanding the child.

*Treatment Decisions and Prognosis.* Decisions about whether to treat and how often the child should be seen are made primarily on the basis of severity and secondarily on the basis of stimulability. Although there is little research on the topic of appropriate treatment decisions, severity appears to have universal acceptance among speech-language pathologists as the most important variable in deciding for or against treatment.

With respect to prognosis, until recently there has been little research about how children normalize (achieve age-appropriate phonology) and how long it takes. However, work by Shriberg, Gruber, and Kwiatkowski (1994) and by Gruber (1999) suggests that some children who receive intervention normalize by about 6 years of age, and that the outer limit for normalization is about 8.5 years. However, the predictors of normalization are not yet known.

—Ann Bosma Smit

**References**


**Further Readings**


**Speech Sampling, Articulation Tests, and Intelligibility in Children with Residual Errors**

Children who have speech sound errors that have persisted past the preschool years are considered to have residual errors (Shriberg, 1994). Typically, these school-age children have substitution and distortion errors rather than deletions, and intelligibility is not usually a primary issue. Children with residual errors generally have acquired the sound system of their language, but they have errors that draw attention to the speaking pattern. The assumption is usually made that they are having difficulty with the articulatory movements needed to produce the acceptable sound and with embedding
that sound into the stream of speech. It should be noted, however, that in the early days of studying communication disorders, authorities such as Van Riper (1978) assumed that the child’s difficulty was first of all perceptual. Until recently, there has been little research about how residual speech sound errors develop, even though some of the earliest research and intervention in communication disorders focused on children with residual errors. The profession has uncovered bits and pieces of information about development after the preschool period, but no coherent picture has emerged to assist in predicting which children will actually make needed changes without intervention. The question is an important one, because if we fail to treat a child at age 6 who is not going to change spontaneously, and instead we wait until age 9, the child has 3 additional years of practice on an error phoneme, and remediation will likely be more difficult. Certain information that can be obtained from speech sampling may provide insight into the prediction question:

- Most residual errors affect a subset of the phonemes of English (“the big 10”): /s z f s t d j θ v r/ (Winitz, 1969). These are typically late-acquired phonemes, but most of them are used correctly by 90% of children by age 8 (Smit et al., 1990).
- The phoneme in error may make a difference. Re-analysis of the Smit et al. (1990) cross-sectional data suggests that children may be less likely to self-correct the alveolar and palatal fricatives and affricates than they are the /r/ and the /θ θ/ (Smit, unpublished).
- The nature and allophonic distribution of the error may make a difference. Stephens, Hoffman, and Daniloff (1986) showed that children with lateral productions of alveolopalatal fricatives and children who substituted back sounds for these fricatives generally did not improve spontaneously, whereas about half of the children with dental errors corrected them. Hoffman, Schuckers, and Daniloff (1980) showed that children who produced the consonantal /r/ allophone correctly some of the time were likely to achieve the other /r/-allophones spontaneously.
- The length of time that the child has made the error may make a difference. It is reasonable to assume that if a child’s production of a phoneme has not changed at all in several years, then spontaneous change is unlikely.
- The child’s developmental history may make a difference. Shriberg (1994) has pointed out that some children had phonological errors as preschoolers, while others did not. On logical grounds, children who had phonological problems earlier are less likely to change without intervention because they have already demonstrated difficulties in learning the sound system of their native language.
- The pattern of change in the child’s errors may be important. Recent research by Gruber (1999) into the time taken for children who are receiving intervention to normalize (achieve age-appropriate phonology) has provided some clues about prognosis. For example, it appears that if the child reduces substitutions and omissions while increasing distortions, that child will take longer to normalize than the child who decreases all types of errors.

Speech Sampling. Speech-language pathologists typically elicit a speech sample using a published test of articulation, supplemented with a conversational speech sample. For a school-age child, the conversational sample should be audio- or video-recorded, with careful attention to the quality of the recording. This sample should include at least 100 different words and 3 minutes of child talking time. If the child has a relatively large number of errors, this speech sample can be transcribed phonetically for further analysis. If the child has just a few sounds in error, then the clinician may decide not to transcribe the entire speech sample. Instead, he tallies all instances of a target phoneme in the sample, determines how many were acceptably produced, and derives a percentage of correctly produced sounds. When these counts are based on a 3-minute sample, this procedure results in a TALK sample (Diedrich, 1971), which is a probe of conversational speech.

Articulation Tests. Some of the first tests that were commercially available in communication disorders were tests of articulation and were designed to assess the development of speech sounds. Typically, tests of this type are based on the single-word naming of pictures without a model from the examiner. Most tests of articulation assess production of all English consonant phonemes in word-initial and word-final position, and possibly English consonant clusters as well. Scoring sheets usually allow explicit comparisons between adult targets and the children’s productions. Most articulation tests result in a summary numerical score that can be compared to normative data. Some currently used tests of articulation include the Templin-Darley Tests of Articulation (Templin and Darley, 1969), the Goldman-Fristoe Test of Articulation—2 (Goldman and Fristoe, 2000), the Smit-Hand Articulation and Phonology Evaluation (SHAPE; Smit and Hand, 1997), and the Photo Articulation Test—Third Edition (Lippke et al., 1987).

Most of the inventory tests do not require that the clinician use narrow transcription, the exception being SHAPE. Rather, broad transcription is generally used, even when the test requires only a notation of “correct,” “substitution,” “omission,” or “distortion.”

The error sounds identified on an inventory test are often examined in light of “ages of acquisition” for those sounds. The age of acquisition is the age at which 75% or 90% of children typically say the sound correctly (Templin, 1957; Smit et al., 1990). The guidelines used by many school districts to determine caseload make reference to these ages of acquisition.

There are other types of tests of articulation besides inventory tests. In particular, there are several tests of contextual variation, among them the McDonald Deep Test of Articulation (McDonald, 1964), the Contextual
pared to the one generated by the examining clinician, the person writes down the child’s words. This gloss is com-
recorded. The examining clinician may want to repeat what she understood the child to say
immediately afterward, so that her gloss is also recorded
articulations of sounds, along with deletions of some
This difficulty may manifest itself as weak or imprecise
One exception is the child who may have a few dis-
Intelligibility.

Intelligibility is defined as a listener’s ability to understand a speaker’s words. Although intel-
ligibility of speech can be reduced in children who have residual errors, the reduction often is not substantial.

Interpreting the Data. Decisions about whether to pro-
vide intervention are often based on multiple factors. These include the child’s age relative to the age of ac-
quisition for the child’s error phonemes, whether or not the child is stimulable for correct production, intelligi-
bility, and the degree to which the child and significant others consider the speech to be a problem. There is little research other than that of Gruber (1999) to go by in establishing prognosis. However, a reasonable assump-
tion is that the older the child who has residual errors, the longer it will take to achieve normalization in a treatment program.

—Ann Bosma Smit

References


Speech Sound Disorders in Children: Description and Classification

Children with speech sound disorders form a heterogeneous group whose problems differ in severity, scope, etiology, course of recovery, and social consequences. Beyond manifest problems with speech production and use, their problems can include reduced intelligibility, risk for broader communication disorders, and academic difficulties, as well as social stigma.

Because of the heterogeneity of children’s speech sound disorders, the description and classification of these disorders have been attempted from a variety of perspectives, with persisting controversy as a predictable result. Nonetheless, one distinction that has garnered relatively universal support is the division of children’s speech disorders into those that are developmental (with onset in early or middle childhood, e.g., before age 9) and those that are nondevelopmental (occurring after that time period and resulting from known causes). Developmental disorders have received substantially more research attention to date.

A second widely accepted distinction separates developmental disorders with known causes from those without. For developmental speech disorders of known causes in children, the terminology has been relatively stable and has typically referenced etiological factors (e.g., speech disorders due to mental retardation, cleft palate). In contrast, the terminology for children’s speech disorders of unknown origin is less stable, reflecting uncertainty about their nature and origin. During the past 30 years, commonly used terms have included functional articulation disorders, phonological disorders (Locke, 1983), articulation and phonological disorders (Bernthal and Bankson, 1998), and persistent sound system disorders (Shelton, 1993).

Proposed classifications of child speech disorders have been advanced along descriptive, predictive, and clinical grounds (Shriberg, 1994). Three classifications currently warrant particular attention either because of empirical support (those associated with Shriberg and Dodd) or practical significance (the Diagnostic and Statistical Manual of Mental Disorders-IV-TR; American Psychiatric Association, 2000).

Currently, the most comprehensive and rigorously studied classification is the Speech Disorders Classification System, developed through a 20-year program of research by Shriberg and his colleagues (Shriberg, 1994, 1999; Shriberg and Kwiatkowski, 1982, 1994a, 1994b, 1994c; Shriberg et al., 1997). This evolving classification is designed to provide a framework for identifying and describing subtypes and testing etiological hypotheses. At this time, it is primarily a research tool.

Within this classification, children’s speech sound disorders of unknown origin are divided into speech delay and residual error categories. Speech delay, with an estimated prevalence of 3.8% among 6-year-olds (Shriberg, Tomblin, and McSweeney, 1999), is characterized by reduced intelligibility and increased risk for broader communication and academic difficulties. It encompasses more severe forms of speech disorder. Residual speech errors, with a tentatively estimated prevalence of 5% among individuals older than age 9 (Shriberg, 1994), is characterized by the presence of at least one speech sound error (often involving distortion of a sibilant fricative or liquid) that persists past the developmental period. Although the category of residual speech errors encompasses less severe forms of speech disorder that are associated with neither reduced intelligibility nor broader communication difficulties, disorders in this category remain of interest for theoretical reasons (i.e., genetic versus environmental origin) and for their potential social and vocational costs, which for some individuals can continue throughout life.

Within the Speech Disorders Classification System, the major categories of speech delay and residual speech errors are further divided according to suspected etiological factors or developmental pattern. Five subtypes of speech delay are postulated in relation to the following possible causes: genetic transmission, early history of recurrent otitis media with effusion (Shriberg et al., 2000), motor speech involvement associated with developmental apraxia of speech (Shriberg, Aram, and Kwiatkowski, 1997), motor speech involvement associated with mild dysarthria, and developmental psychosocial involvement. In each case the etiological factor is considered dominant in a mechanism that is suspected to be multifactorial in nature. Two subtypes of residual speech errors are proposed, those found in association with a documented history of speech delay (residual error-A) and those for which no previous history of speech disorder was reported (residual error-B) (e.g., Shriberg et al., 2001). Ongoing research is aimed at increasing understanding of the causal, developmental, and cognitive processing mechanisms underlying each of these five subtypes.

The classification description described by Dodd (1995) is well motivated from theoretical perspectives.
and based on processing accounts of child speech disorders, but has been less thoroughly validated than the Speech Disorders Classification System. Nonetheless, Dodd’s classification system has been supported by studies that examine characteristics of clinical populations (Dodd, 1995), error patterns across languages (e.g., Fox and Dodd, 2001), bilingual children’s generalization patterns (Holm and Dodd, 2001), and treatment efficacy (Dodd and Bradford, 2000). Thus, its empirical support is growing rapidly. It is intended primarily to be used to aid in differential diagnosis and clinical management, and was proposed as a system that uniquely combined four historical approaches to classifying speech disorders. These four approaches were based on age at onset, severity, causal and maintenance factors, and description of symptoms, respectively.

Dodd’s classification system recognizes five subtypes: articulation disorder, delayed phonological acquisition, consistent deviant disorder, inconsistent disorder, and other. Within this system, an articulation disorder is defined as inability to produce an undistorted version of a speech sound or sounds that are expected, given the child’s age. English sounds that are often affected in such disorders include /s/, /t/, and the interdental fricatives. This label is applied regardless of whether the cause is an anatomical anomaly or is unknown. Delayed phonological acquisition is defined in cases where a child’s speech errors are consistent with those seen in younger, normally developing children. Consistent deviant disorder is the label applied when a child demonstrates a reduced variety of syllable structure use as well as errors that are atypical of those seen in normal development.

Inconsistent disorder is identified when a child’s productions are inconsistent in ways that cannot be explained by complex phonological rules or the effects of linguistic load on production. Ozanne (1995) described a study suggesting that inconsistent disorder represents one subgroup associated with developmental verbal dyspraxia (the condition referred to by Shriberg and others as developmental [or childhood] apraxia of speech). Inconsistent disorder is operationally defined using a 25-item word list. The child is asked to produce each word three times, with inconsistency noted when the child produces at least ten of the words differently on two of the three elicited productions. Dodd’s “other” category encompasses suspected motor speech disorders.

The DSM-IV-TR (American Psychiatric Association, 2000) classification represents the most streamlined classification of children’s speech disorders and one that is perhaps more familiar than others to a broad range of speech pathologists, who use it for billing purposes, and non-speech-language pathologists who come in contact with children with childhood speech disorders. Within this classification, Phonological Disorders 315.39 (formerly Developmental Articulation Disorders) is nested within Communication Disorders. Communication Disorders, in turns, falls under the relatively large category Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence. This category includes, among others, mental retardation, learning disorders (learning disabilities), and pervasive developmental disorders.

In the most recent DSM-IV classification, phonological disorders is defined by the failure to use speech sounds that are expected given the child’s age and dialect. Although subtypes are not described, the American Psychiatric Association’s description of the category phonological disorders acknowledges that errors may reflect difficulties in peripheral production as well as more abstract difficulties in the child’s representation and use of the sound system of the target language. Under comments on differential diagnosis, phonological disorders is described as a possible secondary diagnosis when speech errors in excess of expectations are noted in association with disorders that might be considered known causes for speech difficulties (viz., mental retardation, hearing impairment or other sensory deficit, speech motor deficit, and severe environmental deprivation). Speech difficulties that may be associated with the term “Developmental (or Childhood) Apraxia of Speech” are addressed neither in the DSM-IV criteria nor as a subclassification, although that term is described as a possible label for some forms of phonological disorder in DSM-IV-TR—a revision designed to increase the currency of the DSM without changing the actual classificatory categories.

Classifications of children’s speech disorders are encumbered by demands that they address numerous audiences and unresolved controversies. Among the current audiences to be served are clinicians, researchers, and administrators. Each of the three classifications described here addresses those audiences to a different degree. One of the unresolved controversies that has been addressed to some degree by each is the status of clinically postulated entities such as developmental or childhood apraxia of speech. A second controversy, and one that is being addressed by Shriberg and colleagues, relates to how child speech sound disorders should be conceptualized in relation to other communication disorders that frequently co-occur, but are associated with causal mechanisms that are more ill-defined than those for conditions that fall under child speech disorders with known causes (e.g., hearing loss, cleft palate). Two that are of particular interest are specific language impairment (Shriberg, Tomblin, and McSweeney, 1999) and stuttering (Guitar, 1998). In addition, classifications are ideally consistent with developmental as well as psychological processing accounts of the manifest behaviors associated with child speech disorders. The accounts of Shriberg, Dodd, and their colleagues appear to be pursuing these challenging issues.

—Rebecca J. McCauley

References


Further Readings


Stuttering

Stuttering is a developmental disorder of communication that affects approximately 5% of children born in the United States and Western Europe. Children are at highest risk for beginning to stutter between their second and fourth birthdays. The risk decreases gradually thereafter, with few onsets occurring after 9 or 10 years of age (Andrews and Harris, 1964). The percentage of older children, adolescents, and adults who stutter is much lower, about 0.5%–1.0% (Andrews, 1984; Bloodstein, 1995), and the discrepancy between the percentage of children affected (i.e., incidence) and the percentage of older children and adults who stutter (i.e., prevalence) indicates that 75%–90% of the children who begin to stutter stop. Complete, untreated remissions of stuttering are most likely to occur within 2 years of onset (Andrews and Harris, 1964; Yairi and Ambrose, 1999; Mansson, 2000), with decreasing frequency after that.
Most of the data on the epidemiology of stuttering have been obtained from cross-sectional surveys that asked informants if they or family members currently stutter or had ever stuttered. The credibility of such data is compromised by a number of methodological weaknesses (Yairi, Ambrose, and Cox, 1996). Prospective, longitudinal studies employing trained examiners have been completed in England (Andrews and Harris, 1964) and Denmark (Mansson, 2000). The incidence (5.0%) and remission rate (>75%) reported by both studies were remarkably similar, despite substantial differences in their designs and the populations studied.

The incidence, prevalence, and remission or persistence of stuttering are affected by sex and family histories of stuttering. More than two-thirds of the children who stutter have first-, second-, or third-degree relatives who currently or once stuttered (Ambrose, Yairi, and Cox, 1993). Like most speech-language disorders, stuttering affects more males than females, with about twice as many young male preschoolers affected as females, a ratio that increases to four or more males to every female among adults (Ambrose, Yairi, and Cox, 1997).

Similar ratios have been reported in other countries and cultures (Bloodstein, 1995; Ambrose, Cox, and Yairi, 1997; Mansson, 2000). The increase in male-female ratio with age reflects, in part, higher rates of remission among females (Ambrose, Cox, and Yairi, 1997), whereas family histories of remission and persistence are linked, respectively, to untreated remissions of stuttering within 2 years of onset or its persistence for 3 or more years (Ambrose, Cox, and Yairi, 1997).

Findings from family pedigree studies are consistent with the vertical transmission (i.e., generation to generation) of a genetic susceptibility or predisposition to stutter but are inconsistent with autosomal dominant, recessive, or sex-linked transmissions (Kidd, Heimbuch, and Records, 1981; Yairi, Ambrose, and Cox, 1996). Twin studies have found that stuttering occurs in both members of monozygotic twin pairs much more often than in same-sex, dizygotic twins (e.g., Howie, 1981); however, the lack of concordance of stuttering in some monozygotic twin pairs indicates that both genetic and environmental factors are involved in some, if not all, cases. Segregation analyses suggest that a single major locus is a primary contributor to stuttering phenotypes but that other genes are involved in determining whether or not stuttering persists (Ambrose, Cox, and Yairi, 1997). Research centers in the United States and Europe are currently engaged in linkage analyses designed to identify the specific genes involved.

Current etiological theories reflect diverse beliefs about the nature of stuttering, its origins, and the levels of description that will provide the most useful scientific explanation. However, no theory of origin has achieved general acceptance in the field. There are, for example, cognitive theories, such as Bloodstein’s (1997) anticipatory struggle theory, which hypothesizes that a child’s belief that speech is difficult elicits tension that causes stuttering when he or she tries to speak; psycholinguistic theories (e.g., Ratner, 1997), which propose that the linguistic processes responsible for everyday speech errors are also responsible for developmental stuttering; motor control theories, which link stuttering to sensorimotor or speech motor processes (e.g., Neilson and Neilson, 1987); and multifactor theories, which attribute stuttering to an interplay of the cognitive, linguistic, motor, and affective processes involved in spoken language (e.g., Perkins, Kent, and Curlee, 1991; Smith and Kelly, 1997).

Adult stuttering typically involves various behaviors and affective-cognitive reactions that affect, in varying degrees, interpersonal communication, vocational opportunities, and personal-social adjustment. Frequent repetitions and prolongations of sounds and syllables and blockages of speech occur that cannot be readily controlled (Perkins, 1990). These speech disruptions are often accompanied by facial grimaces and tremors, disrhythmic phonation, and extraneous bodily movements that seem to involve muscle tension and excessive effort.

Stuttering does not occur randomly, but as if constrained by various linguistic variables (Ratner, 1997). For example, it occurs more often than chance would suggest on initial words of phrases, clauses, and sentences, on stressed syllables, and on longer, less familiar words. With the exception of the sentence-initial position, these are the same loci that attract the speech errors of nonstuttering speakers (Fromkin, 1993), suggesting that the same linguistic variables affect both groups’ speech.

The frequency and severity of adults’ stuttering may vary substantially across time, social settings, and contexts, but it often occurs on specific words (e.g., the person’s name) and in situations where stuttering has frequently occurred in the past (Bloodstein, 1995). Thus, adults’ prior stuttering experiences may lead them to avoid specific words and speaking situations and to develop attitudes and beliefs about speaking, stuttering, and themselves that can be more disabling or handicapping than their stuttered speech.

Stuttering is substantially reduced, sometimes eliminated, in a number of verbal activities (Bloodstein, 1995). For example, most adults stutter little or not at all when singing, speaking while alone or with pets or infants, reading or reciting in unison with others, pacing speech with a rhythmic stimulus, and reading or speaking with auditory masking. Stuttering reappears, however, as soon as such activities end.

Much of what is known about adults who stutter has been obtained by studies that compared stuttering and nonstuttering speakers’ motor, sensory, perceptual, and cognitive abilities, as well as the two groups’ affective and personality characteristics. A variety of differences have been found, but usually with substantial overlaps in the data of the two groups. In addition, it is seldom clear if or how such differences might be functionally related to stuttering. A comprehensive review of this work led Bloodstein (1995) to conclude that the two groups are similar, for the most part, except when they speak. Recent advances in brain imaging technology, however, have allowed investigators to compare the brain activity
patterns of the two groups while the subjects read aloud, and the brain activity of stuttering speakers during stuttered and fluent speech.

There are no reliable differences in cerebral blood flow of stuttering and nonstuttering adult men when they are not speaking (Ingham et al., 1996; Braun et al., 1997). A series of H215O positron emission tomography (PET) of the two speaker groups during solo and choral reading conditions found greater right than left hemisphere activity in the supplementary motor and premotor areas (BA 6), anterior insula, and cerebellum and reduced activity in primary auditory areas (BA 41/42) of stuttering speakers in the solo condition, but exactly the opposite pattern of activity in nonstuttering speakers (Ingham, 2001). These differences decreased, however, when fluent speech was induced in stuttering speakers by having them read aloud in unison (i.e., choral reading) with a recording.

A follow-up PET study of subsets of the same two groups was conducted while subjects imagined they were reading aloud (Ingham et al., 2000). Stuttering subjects were instructed to imagine they were stuttering in the solo condition and fluent in the choral condition. The patterns of activity that had occurred when each group read aloud were similar to those observed when speakers merely imagined they were reading. As such studies continue, a much better understanding of the brain-behavior substrates of stuttered and stutter-free or normal speech may be achieved.

Adults who have been stuttering most of their lives often develop various situational fears, social anxieties, lowered expectations, diminished self-esteem, and an array of escape and avoidance behaviors. Prior to initiating treatment, clinicians should obtain a thorough history of an adult’s stuttering and prior treatment, including major current concerns, treatment expectations, and goals, and should assess attitudes, affective reactions and behaviors, and self-concepts that may require treatment. Analyses of samples recorded in various speaking situations document the type, frequency, duration, and overall severity of stuttering. Such information allows clinicians to select appropriate treatment strategies, track progress, and determine when treatment objectives have been achieved. Current treatment strategies focus either on modifying adults’ affective, cognitive, and behavioral reactions to stuttering (e.g., Prins, 1993; Manning, 1996) or on learning speech production techniques (i.e., fluency training) to reduce or eliminate stuttering (e.g., Neilson and Andrews, 1993; Onslow, 1996). Most clinicians apparently prefer a combined strategy to manage the constellation of speech, affective, and cognitive symptoms commonly presented by adults who stutter. No well-controlled clinical trials of stuttering treatments have been reported, but some relapse in treatment gains is common. It is generally agreed, therefore, that complete, permanent recovery from chronic stuttering is rare when stuttering persists into adult life, regardless of the treatment employed. Consequently, local self-help groups of the National Stuttering Association, which promote sharing of experience, information, and support among members, have become an increasingly important component to successful, long-term management of stuttering in adults in the United States.

*See also SPEECH DISFLUENCY AND STUTTERING IN CHILDREN.*

—Richard F. Curlee

**References**


Further Readings


Transsexualism and Sex Reassignment: Speech Differences

According to the Random House Dictionary, a transsexual individual is “A person having a strong desire to assume the physical characteristics and gender role of the opposite sex; a person who has undergone hormone treatment and surgery to attain the physical characteristics of the opposite sex” (Flexner, 1987). Brown and
Rounsley (1996) explain, “Transsexuals are individuals who strongly feel that they are, or ought to be, the opposite sex. The body they were born with does not match their own inner conviction and mental image of who they are or want to be…. This dilemma causes them intense emotional distress and anxiety and often interferes with their day-to-day functioning” (p. 6).

Historically, examples of transsexualism existed in Greek and Roman times, during the Middle Ages, and in the Renaissance (Doctor, 1988). However, the first documented sex reassignment surgery is thought to have been performed around 1923. It involved a man who married at 20 but came to believe he should have been a woman, and took the name of Lili (Hoyer, 1933). The most celebrated case in the United States was that of Christine Jorgensen, who grew up in New York City as a male and had transsexual reassignment surgery performed in Denmark in 1952 (Jorgensen, 1967).

The prevalence of transsexualism is difficult to determine. The United States does not have a national registry to collect information from all possible sources that may deal with the transsexual person. Furthermore, some individuals may remain undiagnosed or may wish to remain “closeted,” or they may travel to foreign countries for sex reassignment surgery.

Data from other countries suggest that one in 30,000 adult males and one in 100,000 adult females seek sex reassignment surgery. However, the overall prevalence figures are greater if one includes transsexuals who do not elect sex reassignment surgery. In the United States, it is estimated that 6000–10,000 transsexuals had undergone sex reassignment surgery by 1988 (Brown and Rounsley, 1996). Spencer (1988) reported that in 1979, more than 4000 U.S. citizens had undergone sex reassignment surgery and about 50,000 were thought to be awaiting surgery. Estimates vary greatly concerning the number of male-to-female transsexuals compared to the number of female-to-male transsexuals, although all estimates indicate that the number of male-to-female transsexuals exceeds the number of female-to-male transsexuals by as much as 3 or 4 to 1 (Oates and Dacakis, 1983; Doctor, 1988; Spencer, 1988; Wolfe et al., 1990).

The transsexual individual (also referred to as a transgendered individual) who seeks therapy and possibly surgery is often referred to a clinic specializing in the treatment of gender dysphoria. Programs that are offered through these clinics include psychological counseling, hormone treatments, and other nonsurgical procedures, which may then lead to the final surgical reassignment surgery. Male-to-female surgery generally requires a single operation and is less expensive than female-to-male surgery, which requires at least three operations, is considerably more expensive, and is not associated with as good aesthetic and functional results as male-to-female surgery (Brown and Rounsley, 1988).

The literature on therapy for the male-to-female transsexual and the female-to-male transsexual represents two extremes. Therapy for the male-to-female transsexual has focused on voice therapy as a major component. Voice therapy for the female-to-male transsexual is virtually nonexistent. In fact, there appears to be considerable agreement among researchers studying the treatment of the transsexual that voice therapy for the female-to-male transsexual is unnecessary because lowering of the fundamental frequency occurs automatically as a result of androgens administered to the female-to-male transsexual (Spencer, 1988; Colton and Casper, 1996). Van Borsel et al. (2000) conducted a two-part study of the voice problems of the female-to-male transsexual. Part 1 was a survey of 16 individuals who had been treated with androgens for at least 1 year by the Gent University Gender Team in Belgium. Questionnaires indicated that 14 of the 16 respondents had experienced a “lower” or “heavier” voice. The remaining two reported that they had a lower-pitched voice before treatment started. Only one of the subjects was not pleased with his voice because of what he perceived as strain in speaking at a lower pitch. Fourteen indicated that voice change was as important as sex reassignment surgery, although 11 of the 16 did not consider the need for speech therapy important. The study confirmed the view that pitch is lowered as a result of androgen treatment and appears to result in an acceptable male voice.

Part 2 was a longitudinal study of the voice change of two female-to-male transsexuals who were administered androgens. Acoustic measures of fundamental frequency, jitter, and shimmer were made of the sustained vowel production of /a/ and the reading of a standard paragraph. The measures for one subject were made over 17 months and for the other subject over 13 months. The results confirmed that the fundamental frequency was substantially reduced for sustained vowel production and reading, although not by more than one octave. Measures of jitter and shimmer were relatively unchanged over time.

The administration of hormones for the male-to-female transsexual has little effect on voice. Some studies have examined the male-to-female transsexual’s changes in fundamental frequency and its relationship to the identification of the voice as a female voice (Bradley et al., 1978; Spencer, 1988; Dacakis, 2000; Gelfer and Schofield, 2000). Although there is some agreement that fundamental frequency is most often perceived as a female voice at 155–160 Hz and above, it is not sufficient alone to identify the male-to-female transsexual as a female speaker (Bradley et al., 1978; Gelfer and Schofield, 2000). Mount and Salmon (1988) conducted a long-range study of a 63-year-old male-to-female transsexual who had undergone sex reassignment surgery. The individual was able to increase her speaking fundamental frequency after 4 months of therapy. However, she was not perceived as a female speaker until formant frequencies had increased, particularly F2 values. This was achieved through the modification of resonance and articulation. Gelfer and Schofield (2000) conducted a study of 15 male-to-female transsexuals with a control group of six biological females and three biological males. All subjects recorded the Rainbow Passage and produced the isolated vowels /a/ and /i/. Twenty undergraduate
psychology majors served as listeners. The only significant differences among subjects were that the "Subjects perceived as female had a higher SFF [speaking fundamental frequency] and a higher upper limit of SFF than subjects perceived as male" (p. 30). Although formant frequencies for /a/ and /i/ were not significantly different between the male-to-female transsexuals perceived as male and those perceived as female, the mean formant frequencies for the perceived female speakers were all higher than those of the transsexual speakers judged to be male. Gunzburger (1995) had six male-to-female transsexual speakers record a list of Dutch words that were also combined into prose. Subjects were asked to read the material in a female manner and a male manner. Acoustic analyses indicated that the central frequency of F3 was systematically higher in the female version. Recordings of two male-to-female speakers that were representative of a male speaker and a female speaker were played to 31 male and female naive listeners, who were asked to identify the sex of the speaker. The perceptual judgments supported the results of the acoustic analyses. It appeared that the male-to-female transsexual speakers judged to be female had F3 formants more like those associated with the female voice. The shorter vocal tract typically found in females produces higher F3 formants than those of males, with a longer vocal tract (Peterson and Barney, 1952; Fant, 1960). Gunzburger (1995) attributed these changes to a decreased vocal cavity length in the perceived female male-to-female transsexual and pointed out that shortening the vocal tract can be accomplished through changes in articulation and retracting the corners of the mouth (p. 347).

According to Stemple, Glaze, and Gerdeman (1995), the male-to-female transsexual not only has to increase her fundamental frequency while being careful not to damage the vocal folds, but also has to learn to modify the resonance, inflection, and intonation to make articulation more precise, and to modify coughing, vocalized pauses, and throat clearing (p. 204).

Therapy for the female-to-male transsexual appears to be less of an issue than therapy for the male-to-female transsexual. Many of the textbooks on voice disorders include a discussion of the therapy needs for the male-to-female transsexual but do not provide any details on procedures, techniques, or concerns for the clinician to consider in the therapy process. De Bruin, Coerts, and Greven (2000) provide the clinician with a detailed approach, including specific goals and subgoals, to follow in therapy for the male-to-female transsexual. Among the major goals are minimizing chest resonance; modifying intonation patterns, articulation, intensity, and rate; and feminizing laughing and coughing. In addition, they address other verbal and nonverbal aspects of feminine communication such as gestures, movements, greetings, shaking hands, dress, and hairdo. The authors briefly discuss laryngeal surgery but conclude that it only results in raising the fundamental frequency (which is not in itself sufficient to guarantee a feminine voice) and that the results of this surgery are not predictable. Batin (1983) includes vocabulary and language forms and uses videotapes of the male-to-female transsexual to teach the individual how to walk, sit down, and enter a room. Chaloner (1991) provides case histories, and uses role playing in group therapy to help the male-to-female transsexual become more successful in "living the female role" (p. 330). Future research on the assessment and treatment of transgendered individuals should provide the clinician with a larger repertoire of approaches to assist the transsexual individual in making the transition to a different sexual role.

—John M. Pettit

References


**Further Readings**


---

**Ventilator-Supported Speech Production**

When breathing becomes difficult or impossible, it may be necessary to use a ventilator to sustain life. Usually the need for a ventilator is temporary, such as during a surgical procedure. However, if breathing difficulty is chronic, ventilatory support may be required for an extended period, sometimes a lifetime. The main indications for ventilatory support are respiratory insufficiency resulting in hypventilation (not enough gas moving into and out of the lungs), hypoxemia (not enough oxygen in the arterial blood), or hypercapnea (too much carbon dioxide in the blood). Many medical conditions can cause severe respiratory insufficiency requiring ventilatory support. Examples include cervical spinal cord injury (rostral enough to impair diaphragm function), muscular dystrophy, amyotrophic lateral sclerosis, and chronic obstructive pulmonary disease.

Several types of ventilator systems are available for individuals with respiratory insufficiency, including positive-pressure ventilators, negative-pressure ventilators, phrenic nerve pacers, abdominal pneumobelts, and rocking beds (Banner, Blanch, and Desautels, 1990; Hill, 1994; Levine and Henson, 1994). Positive-pressure ventilators operate by “pushing” air into the pulmonary system for inspiration, whereas negative-pressure ventilators work to lower the pressure around the respiratory system and expand it for inspiration. Phrenic nerve pacers stimulate the phrenic nerves and cause the diaphragm to contract to generate inspiration. Abdominal pneumobelts displace the abdomen inward (by inflation of a bladder) to push air out of the pulmonary system for expiration, and then allow the abdomen to return to its resting position (by deflation of the bladder) for inspiration. Rocking beds are designed to move an individual upward toward standing and downward toward supine to drive inspiration and expiration, respectively, using gravitational force to displace the abdomen and diaphragm. All of these systems are currently used (Make et al., 1998); however, the most commonly used one today and the one that the speech-language pathologist is most likely to encounter in clinical practice is the positive-pressure ventilator (Spearman and Sanders, 1990).

The positive-pressure ventilator uses a positive-pressure pump to drive air through a tube into the pulmonary system. The tube can be routed through (1) the larynx (in this case, it is called an endotracheal tube), such as during surgery or acute respiratory failure; (2) the upper airway, via a nose mask, face mask, or mouthpiece (this is called noninvasive ventilation); or (3) a tracheostoma (a surgically fashioned entry through the anterior neck to the tracheal airway). The latter two modes of delivery are used in individuals who need long-term ventilatory support. With noninvasive positive-pressure ventilation, speech is produced in a relatively normal manner. That is, after inspiratory air from the ventilator flows into the nose and/or mouth, expiration begins and speech can be produced until the next inspiration is delivered. The situation is quite different, however, when inspiratory air is delivered via a tracheostoma. In some cases it is not possible to produce speech with the ventilator-delivered air because the air is not allowed to reach the larynx. This occurs when the tracheostomy tube, which is secured in the tracheostoma and provides a connection to the ventilator’s tubing, is configured so as to block airflow to the larynx. This is done by inflating a small cuff that surrounds the tube where it lies within the trachea. However, if the cuff is deflated (or if there is no cuff), it is possible to speak using the ventilator-delivered air. Because the air from the ventilator enters below the larynx, speech can be
produced during both the inspiratory and the expiratory phase of the ventilator cycle (Fig. 1). During the inspiratory phase, speech production competes with ventilation because the ventilator-delivered air that flows through the larynx to produce speech is routed away from the pulmonary system, where gas exchange takes place (i.e., oxygen is exchanged for carbon dioxide). For this and other reasons, the act of speaking with a tracheostomy and positive-pressure ventilator is challenging, and the resultant speech often is quite abnormal.

Positive-pressure ventilators can be adjusted to meet each individual's ventilatory needs. These adjustments typically are determined by the pulmonologist and executed by the respiratory therapist. The most basic adjustments involve setting the tidal volume and breathing rate, the product of which is the minute ventilation (the amount of air moved into or out of the pulmonary system per minute). These parameters are adjusted primarily according to the client's body size and breathing comfort, and their appropriateness is confirmed by blood gas measurements. Most ventilators allow adjustment of other parameters, such as inspiratory duration, magnitude and pattern of inspiratory flow, fraction of inspired oxygen, and pressure at end-expiration (called positive end-expiratory pressure, or PEEP), among others. How these parameters are adjusted influences ventilation and can also have a substantial influence on speech production.

The speech produced with a tracheostomy and positive-pressure ventilator is usually abnormal. Some of its common features are short utterances, long pauses, and variable loudness and voice quality (Hoit, Shea, and Banzett, 1994). The mechanisms underlying these speech features are most easily explained by relating them to the tracheal pressure waveform associated with ventilator-supported speech. This waveform is shown schematically in Figure 2, along with a waveform associated with normal speech production. Whereas tracheal pressure during normal speech production is positive (i.e., above atmospheric pressure), generally low in amplitude (i.e., in the range of 5–10 cm H2O), and relatively unchanging throughout the expiratory phase of the breathing cycle, tracheal pressure during ventilator-supported speech production is generally fast-changing (i.e., rapidly rising during the inspiratory phase of the ventilator cycle and rapidly falling during the expiratory phase of the cycle), high-peaked (approximately 35 cm H2O in the figure), and not always above atmospheric pressure (i.e., during the latter 2 s of the cycle in the figure). These waveforms can also be examined relative to the minimum pressure required to maintain vibration of the vocal folds for phonation (labeled Threshold Pressure in the figure). From this comparison, it is clear that the tracheal pressure associated with normal speech production exceeds this threshold pressure throughout the cycle (expiratory phase), whereas the pressure associated with the ventilator-supported speech production is below the threshold pressure for nearly half the cycle. This latter observation largely explains why ventilator-supported speech is characterized by short utterances and long pauses. The periods during which the pressure is above the voicing threshold pressure are relatively short (compared with normal speech-related expirations) and the periods during which the pressure is below that threshold are relatively long (compared with normal speech-related inspirations). The reason why ventilator-supported speech is variable in loudness and voice quality has to do with the fast-changing nature of the tracheal pressure waveform. The rapid rate at which the pressure rises and falls makes it impossible for the larynx to make the adjustments necessary to produce a steady voice loudness and quality.

There are several strategies for improving ventilator-supported speech. One set of strategies is mechanical in nature and involves modifying the tracheal pressure waveform. Specifically, speech can be improved if the tracheal pressure stays above the voicing threshold.
pressure for a longer portion of the ventilator cycle (to increase utterance duration and decrease pause duration) and if it changes less rapidly and does not peak as highly (to decrease variability of loudness and voice quality). The tracheal pressure waveform can be modified by adjusting certain parameters on the ventilator (such as those mentioned earlier) or by adding external valves to the ventilator system (e.g., Dikeman and Kazandjian, 1995; Hoit and Banzett, 1997). Ventilator-supported speech can also be improved using behavioral strategies. Such strategies include the use of linguistic manipulations designed to hold the floor during conversation (e.g., breaking for obligatory pauses at linguistically inappropriate junctures) and the incorporation of another sound source to supplement the laryngeal voicing source (e.g., buccal or pharyngeal speech).

Evaluation and management of the speech of a client with a tracheostomy and positive-pressure ventilator involves a team approach, with the team usually consisting of a speech-language pathologist working in collaboration with a pulmonologist and a respiratory therapist. Such collaboration is critical because speech production and ventilation are highly interdependent in a client who uses a ventilator. An intervention designed to improve speech will almost certainly influence ventilation, and an adjustment to ventilation will most likely alter the quality of the speech. As an example, a speech-language pathologist might request that a client be allowed to deflate his cuff so that he can speak. Cuff deflation should not compromise ventilation as long as tidal volume is increased appropriately (Bach and Alba, 1990). By understanding the interactions between speech production and ventilation, clinicians can implement interventions that optimize spoken communication without compromising ventilation, thereby improving the overall quality of life in clients who use ventilators.

—Jeanette D. Hoit

References


Further Readings


Part III: Language
Agrammatism

Agrammatism is a disorder that leads to difficulties with sentences. These difficulties can relate both to the correct comprehension and the correct production of sentences. That these difficulties occur at the sentence level is evident from the fact that word comprehension and production can be relatively spared.

Agrammatism occurs in many clinical populations. In patients with Wernicke’s aphasia, for instance, agrammatism has been established both for comprehension (Lukatela, Schankweiler, and Crain, 1995) and for production (Haarmann and Kolk, 1992). Agrammatic comprehension has been demonstrated in patients with Parkinson’s disease (Grossman et al., 2000), Alzheimer’s disease (Waters, Caplan, and Rochon, 1995), and in children with specific language disorders (Van-der-Lely and Dewart, 1986). Agrammatic comprehension has even been demonstrated in normal subjects processing under stressfull conditions (Dick et al., 2001). However, agrammatism has been studied most systematically in patients with Broca’s aphasia, and it is this group this review will focus on.

Symptoms of agrammatic comprehension are typically assessed by presenting a sentence to the subject and asking the subject to pick from a number of pictures the one depicting the proper interpretation of the sentence. Another procedure is to ask subjects to act out the meaning of the sentence with the help of toy figures. The main symptoms thus established are the following: (1) Sentences in which the two thematic roles can be reversed (e.g., “The cat is chasing the dog”) are substantially harder to understand than their nonreversible counterparts (“The cat is drinking milk”) (Caramazza and Zurif, 1976; Kolk and Friederici, 1985). Roughly speaking, thematic roles specify who is doing what to whom. (2) Sentences with noncanonical ordering of thematic roles around the verb are harder to comprehend than ones with canonical ordering. In English, the order of the active sentence is considered to be canonical: agent-action-patient (or subject-verb-object). Sentences with a word order deviating from this pattern are relatively difficult to understand. Thus, passive constructions are harder to understand than active ones (Schwartz, Saffran, and Marin, 1980; Kolk and van Grunsven, 1985), and object relative sentences (“The boy whom the girl pushed was tall”) are harder than subject relative sentences (“The boy who pushed the girl was tall”) (Lukatela, Schankweiler, and Crain, 1995; Grodzinsky, 1999), to mention the most frequently studied contrasts. (3) Sentences with a complex—more deeply branched—phrase structure are harder to understand than their simple counterparts, even if they have canonical word order. For instance, a locative construction (e.g., “The letter is on the book”) is harder to understand than a simple active construction (“The sailor is kissing the girl”), even if subjects are able to comprehend the locative proposition as such (Schwartz, Saffran, and Marin, 1980; Kolk and van Grunsven, 1985). Furthermore, sentences with embedded clauses (“The man greeted by his wife was smoking a pipe”) are harder to comprehend than sentences with two conjoined sentences (“The man was greeted by his wife and he was smoking a pipe”) (Goodglass et al., 1979; Caplan and Hildebrandt, 1988).

Agrammatic production has attracted much less attention than agrammatic comprehension. Symptoms of agrammatic production have traditionally been assessed by analyzing spontaneous speech (Goodglass and Kaplan, 1983; Rochon, Saffran, Berndt, and Schwartz, 2000). Four types of symptoms of spontaneous speech have been established. (1) Reduced variety of grammatical form. If sentences are produced at all, they have little subordination or phrasal elaboration. (2) Omission of function words (articles, pronouns, auxiliaries, prepositions, and the like) and inflections. (3) Omission of main verbs. (4) A slow rate of speech. Whereas these symptoms have been established in English-speaking subjects, similar symptoms occur in many other languages (Menn and Obler, 1990). A number of studies have attempted to elicit production of grammatical morphology and word order in agrammatic patients. A complicating factor is that there are systematic differences between spontaneous speech and elicited speech. In particular, function word omission is less frequent in elicited speech and function word substitution is more frequent (Hofstede and Kolk, 1994). The following symptoms have been observed on elicitation tests. (1) Grammatical word order is impaired (Saffran, Schwartz, and Marin, 1980). (2) It is more impaired in embedded clauses than in main clauses (Kolk and van Grunsven, 1985). (3) Inflection for tense is harder than inflection for agreement (Friedmann and Grodzinsky, 1997). (4) Sentences with noncanonical ordering of thematic roles appear harder to produce than their canonical counterparts (Caplan and Hanna, 1998; Bastiaanse and van Zonneveld, 1998; but see also Kolk and van Grunsven, 1985).

The localization of agrammatism is variable. With respect to both production and comprehension, agrammatism is associated with lesions across the entire left perisylvian cortex.

Theories of agrammatism abound. Some researchers claim that differences between patients are so great that a unitary theory will not be possible (Miceli et al., 1989). Extant theories pertain either to comprehension or to production. This is justified by the fact that agrammatic production and comprehension can be dissociated (Miceli et al., 1983). The most important approaches are the following. The TRACE DELETION HYPOTHESIS about agrammatic comprehension holds that traces, or empty elements resulting from movement transformations according to generative linguistic theories, are lacking (Grodzinsky, 2000). The mapping hypothesis maintains that it is not a defect in the structural representation that is responsible for these difficulties but a defect in the procedures by which these representations are employed to derive thematic roles (Linebarger, Schwartz, and Saffran, 1983). Finally, a number of hypotheses claim a processing limitation to be the bottleneck. The limitation may relate to working memory capacity (Caplan
and Waters, 1999), altered weights or increased noise in a distributed neural network (Dick et al., 2001), or a slowdown in syntactic processing (Kolk and van Grunsven, 1985). With respect to production, the tree truncation hypothesis maintains that damage to a particular node in the syntactic tree leads to the impossibility of processing any structure higher than the damaged node (Friedmann and Grodzinsky, 1997). Finally, the adaptation theory of agrammatic speech (Kolk and van Grunsven, 1985) maintains that the underlying deficit is a slowing down of the syntactic processor. A second claim is that the actual slow, telegraphic output results from the way patients adapt to this deficit.

Treatment programs for agrammatism vary from theoretically neutral syntax training programs (Helms-estabrooks, Fitzpatrick, and Barresi, 1981), to programs motivated by the mapping hypothesis (Schwartz et al., 1994) or by the trace deletion hypothesis (Thompson et al., 1996). The reduced syntax therapy proposed by Springer and Huber (2000) takes a compensatory approach to treatment and fits well with the adaptation theory.

—Herman Kolk

References


Further Readings


**Agraphia**

Agraphia (or dysgraphia) is the term used to describe an acquired impairment of writing. The impairment may result from damage to any of the cognitive, linguistic, or sensorimotor processes that normally support the ability to spell and write. These procedures can be conceptualized within the framework of a cognitive model of language processing such as that shown in Figure 1 (Ellis, 1988; Shallice, 1988; Rapcsak and Beeson, 2000). According to the model, the writing process can be divided into central and peripheral components. The central components are linguistic in nature and are responsible for the retrieval of appropriate words and provision of information about their correct spelling. Peripheral procedures serve to translate spelling knowledge into handwriting, and to guide the motor control for appropriate movements of the hand.

When the system is working normally and an individual wants to write a familiar word, the relevant concepts in the semantic system activate representations in the memory store for learned spellings (i.e., the orthographic output lexicon). Access to this lexicon via the semantic system is referred to as the lexical-semantic spelling route. In contrast, when the individual attempts to spell unfamiliar words or pronounceable nonwords (such as *flig*), reliance on knowledge of sound-to-letter correspondences allows the assembly of plausible spellings by a process referred to as phoneme-grapheme conversion. This alternative means of spelling is depicted in Figure 1 by the arrow from the phonological buffer (where phonological information is held) to the graphemic buffer (where the assembled spelling is held). Spelling in this manner is considered a nonlexical process, because spellings are not retrieved as whole words from the lexicon. Spellings generated by the lexical-semantic and nonlexical spelling routes are subsequently processed in the graphemic buffer. This buffer serves as an interface between central spelling processes and the
peripheral procedures that support the production of handwriting. Peripheral writing procedures are accomplished through a series of hierarchically organized stages that include letter selection (referred to as allographic conversion), motor programming, and the generation of graphic innervatory patterns.

Clinical assessment of spelling and writing provides an understanding of the nature and degree of impairment to specific processes as well as the availability of residual abilities (Kay, Lesser, and Coltheart, 1992; Beeson and Hillis, 2001; Beeson and Rapcsak, 2002). Damage to specific components of the spelling process may result in identifiable agraphia syndromes with relatively predictable lesion sites (Roeltgen, 1993, 1994; Rapcsak and Beeson, 2000, 2002). Central agraphia syndromes reflect damage to the lexical-semantic or nonlexical spelling routes, or the graphemic buffer, and result in similar impairments across different modalities of output (e.g., written spelling, oral spelling, typing). Central agraphia syndromes include lexical agraphia, phonological agraphia, deep agraphia, and graphemic buffer agraphia. Peripheral agraphia syndromes reflect damage to writing processes that are distal to the graphemic buffer. Dysfunction primarily affects the selection or production of letters in handwriting. These syndromes include allographic disorders, apraxic agraphia, and nonapraxic disorders of neuromuscular execution (Roeltgen, 1993; Rapcsak, 1997; Rapcsak and Beeson, 2000). An individual may have impairment to multiple components of the writing process so that the agraphia profile does not conform to a recognized syndrome.

Lexical agraphia (also called surface agraphia) is a central agraphia syndrome that results from damage to the lexical-semantic spelling route. It is characterized by the loss or unavailability of orthographic knowledge, so that spelling is accomplished by phoneme-grapheme conversion and words are spelled as they sound. Spelling accuracy is strongly influenced by orthographic regularity in that regular words (e.g., bake) and nonwords are spelled correctly, but attempts to spell words with irregular sound-to-spelling relationships result in phonologically plausible errors (e.g., cough—coff). Low-frequency, irregular words are especially vulnerable to error. In addition, if the semantic influence on spelling is impaired, there is difficulty writing homophonic words that cannot be spelled correctly without reference to meaning (e.g., dear—deer). Lexical agraphia is typically seen following damage to left extrasylvian temporoparietal regions. The syndrome has also been described in patients with Alzheimer’s disease and in semantic dementia.

Phonological agraphia and deep agraphia are central agraphia syndromes attributable to dysfunction of the nonlexical spelling route. In both syndromes, spelling is accomplished primarily via a lexical-semantic strategy, and individuals have difficulty using phoneme-grapheme conversion to spell unfamiliar words or nonwords. In phonological agraphia, the spelling of familiar words, both regular and irregular, may be relatively spared;
however, in deep agraphia, there is concomitant impairment of the lexical-semantic spelling route, so that semantic errors are prevalent (e.g., boy—girl). In both syndromes, spelling accuracy is better for highly frequent, concrete words (e.g., house) than for low-frequency, abstract words (e.g., honor). There is also an influence of grammatical word class in that nouns are easier to spell than function words such as prepositions, pronouns, and articles. Other spelling errors may include morphological errors (talked—talking) and functor substitutions (as—with). As in any of the central agraphia syndromes, patients may recall only some of the letters of the target word, reflecting partial orthographic knowledge. Phonological and deep agraphia are associated with damage to the perisylvian language areas, including Broca’s area, Wernicke’s area, and the supramarginal gyrus. Deep agraphia in patients with extensive left hemisphere lesions may reflect reliance on the right hemisphere for writing (Rapcsak, Beeson, and Rubens, 1991).

Graphemic buffer agraphia reflects impairment of the ability to retain orthographic representations in short-term memory as the appropriate graphic motor programs are selected and implemented. Damage to the graphemic buffer leads to abnormally rapid decay of information relevant to the order and identity of stored graphemes. Spelling accuracy is notably affected by word length because each additional grapheme increases the demand on limited storage capacity. In contrast to other central agraphia syndromes, spelling in graphemic buffer agraphia is not significantly influenced by lexical status (words versus nonwords), lexical-semantic features (frequency, concreteness, grammatical class), or orthographic regularity. Characteristic spelling errors include letter substitutions, additions, deletions, and transpositions (e.g., garden—garned). These errors are observed in all spelling tasks and across all modalities of output (writing, oral spelling, typing). Lesion sites in patients with graphemic buffer agraphia have been variable, but left parietal and frontal cortical involvement is common.

Allographic disorders are peripheral writing impairments that reflect the breakdown of procedures by which orthographic representations are mapped to letter-specific graphic motor programs. Allographic disorders are characterized by an inability to activate or select appropriate letter shapes, whereas oral spelling is preserved. Patients may have difficulty that is specific to writing upper- or lowercase letters, or they may produce case-mixing errors (e.g., pApeR). Other patients produce well-formed letter substitution errors that bear physical similarity to the target. Allographic disorders are usually associated with damage to the left parieto-occipital region.

Apraxia agraphia is a peripheral writing impairment caused by damage to graphic motor programs, or it may reflect an inability to translate information contained in these programs into specific motor commands. Apraxic agraphia is characterized by poor letter formation that cannot be attributed to sensorimotor impairment (i.e., weakness, deafferentation) or damage to the basal ganglia (i.e., tremor, rigidity) or cerebellum (i.e., ataxia, dysmetria). Typical errors of letter morphology include spatial distortions and stroke omissions or additions, which may result in illegible handwriting. Spelling by other modalities (e.g., oral spelling) is typically spared. In right-handers, apraxic agraphia is associated with damage to a left hemisphere cortical network dedicated to the motor programming of handwriting movements. The major functional components of this neural network include posterior-superior parietal cortex (including the region of the intraparietal sulcus), dorsolateral premotor cortex, and the supplementary motor area (SMA). Cerebral lesions in right-handers may be accompanied by unilateral apraxic agraphia of the left hand.

Writing disorders attributable to impaired neuromuscular execution are caused by damage to motor systems involved in generating graphic innervatory patterns. Poor motor control results in defective control of writing force, speed, and amplitude. Such writing disorders reflect the specific underlying disease or locus of damage. In the case of Parkinson’s disease, micrographia results from reduced force and amplitude of movements of the hand. In patients with cerebellar dysfunction, movements of the pen may be disjointed and erratic. Breakdown of graphomotor control in these neurological conditions suggests that the basal ganglia and the cerebellum, working in concert with dorsolateral premotor cortex and the SMA, are critically involved in the selection and implementation of kinematic parameters for writing movements. Obviously, patients with hemiparesis often have weakness and spasticity of the hand and limb that markedly impairs their ability to write with the preferred hand.

Behavioral treatments for agraphia may target central or peripheral components of the writing process (Behrmann and Byng, 1992; Carlomagno, Iavarone, and Colombo, 1994; Hillis and Caramazza, 1994; Patterson, 1994; Beeson and Hillis, 2001; Beeson and Rapcsak, 2002). Treatments for central agraphias may be directed toward the lexical-semantic or nonlexical spelling procedures. In contrast, treatments for peripheral agraphias are designed to improve the selection and implementation of graphic motor programs for writing. In general, agraphia treatments are designed to strengthen damaged processes and take advantage of residual abilities.

See also alexia; phonological analysis of language disorders in aphasia; phonology and adult aphasia.

—Pelagie M. Beeson and Steven Z. Rapcsak

References


Alexia

Alexia, or acquired impairment of reading, is extremely common after stroke, dementia, or traumatic brain injury. Reading can be affected in a variety of different ways, leading to a number of different clinical syndromes or types of alexia. To understand these alexic syndromes, it is necessary to appreciate the cognitive processes underlying the task of reading words. Reading aloud a familiar word, such as _leopard_, normally entails at the very least seeing and perceiving the entire written letter string, recognizing the word as a known word (by accessing a stored representation of the learned spelling of the word, in the orthographic input lexicon), accessing its meaning (or semantic representation), and accessing the pronunciation (the stored sound of the word, in the phonological output lexicon), as well as activating motor speech mechanisms for articulating the word. Even the first of these components, seeing and perceiving the entire written letter string, requires complex visual-perceptual skills, including computation of several levels of spatial representation, before the stored representations can be accessed. Furthermore, reading an unfamiliar word—say, an unfamiliar surname—entails access to print-to-sound conversion, or grapheme-to-phoneme conversion (GPC) mechanisms. Familiar words can also be read via GPC mechanisms, but the accuracy of pronunciation will depend on the “regularity” of the word—the extent to which the word conforms to typical GPC rules. For example, _sail_ but not _yacht_ can be read accurately via GPC mechanisms.

These components underlying the reading process are schematically depicted in Figure 1 (see Hillis and Caramazza, 1992, or Hillis, 2002, for a review of the evidence for various components of this model). Various features of this model are controversial, such as the precise nature and arrangement of the components, the degree to which they interact, and whether the various levels of representation are accessed in parallel or serially (Shallice, 1988; Hillis and Caramazza, 1991; Plaut and Shallice, 1993; Hillis, 2002). Nevertheless, most models of naming include most of the components depicted in Figure 1. Neurological impairment can selectively impair any one or more of these components of the reading process, with different consequences in terms of the types of errors produced and the types of stimuli that are affected. In addition, because several of the components of reading are cognitive mechanisms that are also involved in other tasks, damage to one of these shared components will have predictable consequences for reading and other tasks. For example, impairment at the level of semantics, or word meaning, will affect not only the reading of familiar irregular words but also the naming and comprehension of words. Thus, it is possible to identify what component of the reading system is impaired by

![Figure 1. The cognitive processes underlying reading.](image-url)
considering the types of errors produced by the individual, the types of words that elicit errors, and the accuracy of performance across other language tasks, such as spoken naming and comprehension. The consequences of damage to each level of the reading process are discussed here.

**Visual Attention and Perception.** To read a familiar or unfamiliar word, the string of letters must be accurately perceived in the correct order and converted to a series of graphemes (abstract letter identities, without a particular case or font). Perception of the printed word can break down when there is (1) poor visual acuity or visual field cut, (2) impairment in distinguishing individual letters or symbols in a string (attentional dyslexia; Shallice, 1988), or (3) impairment in perception of more than one object or feature at a time (called simultagnosia; Parkin, 1996). An individual with simultagnosia might read the word *chair* as “h” or “i”. Finally, individuals with damage to the nondominant (usually right) hemisphere of the brain often fail to perceive the side of a visual stimulus like a word that is contralateral to the brain damage. Such an impairment, known as neglect dyslexia, results in reading errors such as *chair* → “fair,” *spool* → “pool,” and *love* → “glove,” errors that entail substitution, deletion, or insertion of letters on the left side (or initial letters) of words (Kinsbourne and Warrington, 1962; see papers in Riddoch, 1991, for reviews). Depending on the level of spatial representation affected, reading accuracy is sometimes improved by moving the printed word to the unaffected side of space or by spelling the word aloud to the person (see Hillis and Caramazza, 1995, for a discussion of various types of neglect dyslexia resulting from damage to distinct levels of spatial representation that are computed prior to accessing a stored graphemic representation). All types of reading stimuli are likely to be affected in neglect dyslexia, although words that have final letter strings in common with other words often elicit the most errors. For example, the words *light*, *fight*, *might*, *right*, *tight*, *sight*, *blight*, *bright*, *slight*, and so on are all likely to be read as the same word, since only the final letters are perceived and used to access the stored graphemic representation for recognition. Pseudo-words also elicit comparable errors (e.g., *glamp* → “lamp” or “damp”). An individual with impairment in computing one or more levels of spatial representation will usually also make errors in perceiving the left side of nonlinguistic visual stimuli (Hillis and Caramazza, 1995), although exceptional cases of pure neglect dyslexia, without other features of hemispatial neglect, have been reported (Costello and Warrington, 1987; Patterson and Wilson, 1990).

**Orthographic Input Lexicon.** Impairment at the level of accessing learned spellings of familiar words, or stored graphemic representations that constitute the “orthographic input lexicon,” results in impaired recognition of written words despite accurate perception of the letters. The individual with this impairment will often fail to distinguish familiar from unfamiliar words, or pseudowords (e.g., *glamp*), in a task known as *lexical decision*. Sometimes such an individual will read each letter in the string aloud serially, which seems to facilitate access to the orthographic input lexicon (resulting in letter-by-letter reading; see papers in Coltheart, 1998). If GPC mechanisms are intact, these mechanisms may be used to read even familiar words, resulting in “regularization” of irregular words (e.g., *one* → “own”). Other errors are predominantly visually similar words (e.g., *though* → “touch”). Oral reading of all types of words may be affected, although very familiar words—those frequently encountered in reading—may be relatively spared. Since the orthographic input lexicon is not involved in other linguistic tasks, damage to this cognitive process does not cause errors in other tasks. Therefore, individuals with impairment of this component are said to have pure alexia.

**Semantic System.** Disruption of semantic representations is often incomplete, such that the meanings that are accessed are often impoverished, and only certain categories of words are affected. For example, an alexic patient may read *dog* as “eat” if an incomplete semantic representation of *dog* is accessed that specifies only *animal,* or *mammal,* *domesticated,* *quadraped,* etc., without information about what differentiates a dog from a cat. Thus, most errors are semantically related words, such as *robin* → “cardinal” or *robin* → “bird” (errors called semantic paralexias). However, if GPC mechanisms are available, these may be used to read all types of words, or used to block semantic paralexias. Or GPC mechanisms may be combined with partial semantic information to access the correct phonological representation in the output lexicon, so that the individual can read aloud words better than he or she can understand words (Hillis and Caramazza, 1991). Often there is especially incomplete semantic information to distinguish abstract words, so that abstract words are read less accurately than concrete words. Similarly, functors are read least accurately, often with one functor substituted for another (e.g., *therefore* → “because”); verbs are read less accurately than adjectives; and adjectives are read less accurately than nouns (Coltheart, Patterson, and Marshall, 1980).

If GPC mechanisms are also impaired (a commonly occurring deficit), pseudo-words and unfamiliar words cannot be read. Since the semantic system is shared by the tasks of naming and comprehension, semantic errors are also made in oral and written naming and in comprehension of spoken and written words (Hillis et al., 1990).

**Phonological Output Lexicon.** Impairment in accessing phonological representations for output results in poor oral reading despite accurate comprehension of printed words. For example, *gray* might be read as “blue” but defined as “the color of hair when you get old” (from Caramazza and Hillis, 1990). Again, if GPC mechanisms are available, these mechanisms may be used to
read aloud all types of words, resulting in regularization errors on irregular words. Otherwise, errors may be semantically related (e.g., *fork* → “spoon”), or phonologically related to the target (e.g., *choir* → “quar*”). Phonological representations of words that are used more frequently may be more accessible than other words, so that high-frequency words are read more accurately than low-frequency words. Since the phonological output lexicon is also essential for oral naming and spontaneous speech, the person will make similar errors on these tasks as in reading.

**Types of Alexia**

A number of alexic syndromes consisting of a particular pattern of frequently co-occurring symptoms in reading have been described. (These types of alexia are also known as acquired dyslexia.) Individuals with a particular alexic syndrome may have different underlying deficits, however. To identify which component of the reading process is impaired it is necessary to know not only the error types and the types of words that are misread, but also the individual’s pattern of performance on other lexical tasks, such as naming and comprehension.

Surface alexia or surface dyslexia refers to a pattern of reading that reflects use of GPC mechanisms to read both familiar and unfamiliar words, so that irregular words are often read as regularization errors (e.g., *bear* → “beer”; see papers in Patterson, Coltheart, and Marshall, 1985). Regular words, which can be read accurately via GPC mechanisms, are more likely to be read correctly than irregular words. Comprehension of homophones, such as *eight* and *ate*, may be confused. Oral reading of pseudo-words is accurate. This pattern can be seen with damage to any level of the reading system that requires reliance on GPC mechanisms to bypass the damaged component. For example, Table 1 describes features of three patients with surface alexia who each had impairment at different levels of lexical representation but had intact GPC mechanisms.

*Phonological alexia or phonological dyslexia refers to impairment in use of GPC mechanisms, so that the reader is unable to compose a plausible pronunciation of unfamiliar words or pseudo-words. In addition, occasional semantic paralexias or functor substitutions are produced, presumably because these errors are not blocked by GPC mechanisms (see Beauvois and Due-ouesne, 1979; Goodglass and Budin, 1988; Shallice, 1988).*

Deep alexia or deep dyslexia is a pattern that arises when there is damage to both GPC mechanisms and another component of the “semantic route” of reading: the semantic system and/or the phonological output lexicon (see Coltheart, Patterson, and Marshall, 1980). Semantic paralexias and functor substitutions are invariably produced, although visually similar word errors and derivational errors (e.g., *write* → “writer”; *predicted* → “prediction”) are also common. Concrete words are read more accurately than abstract words, and there is the following grammatical category effect: nouns > adjectives > verbs > functors (nouns most accurate). Table 2 characterizes patterns of performance across tasks in patients with deep dyslexia with damage to different components of the semantic route. It has been argued that the pattern of reading errors reflects reliance on the nondominant hemisphere’s rudimentary language capabilities (Coltheart, 1980), although direct evidence for this proposal is lacking.

N*egot dyslexia, attentional dyslexia, and pure alexia were described under impairments of specific components.*

Neurological disease or focal brain damage can disrupt one or more relatively distinct cognitive mechanisms that underlie the task of reading, resulting in different patterns of reading impairment. It is generally possible to identify the impaired components by determining the types of errors made, the types of stimuli that

---

**Table 1. Characteristics of Reported Individuals with Surface Alexia (Compensatory Use of GPC Mechanisms)**

<table>
<thead>
<tr>
<th>Error types (example of errors)</th>
<th>Lexical decision</th>
<th>Lexical decision</th>
<th>Lexical decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regularization (<em>bear</em> → “beer”)</td>
<td>Impaired</td>
<td>Impaired</td>
<td>Impaired</td>
</tr>
<tr>
<td>Written word comprehension</td>
<td>* regularly read as “regularize”*</td>
<td>* regularly read as “regularize”*</td>
<td>* regularly read as “regularize”*</td>
</tr>
<tr>
<td>Oral naming</td>
<td>* Impaired</td>
<td>* Impaired</td>
<td>* Impaired</td>
</tr>
<tr>
<td>GPC mechanisms</td>
<td>* Intact</td>
<td>* Intact</td>
<td>* Intact</td>
</tr>
<tr>
<td>Level of deficit</td>
<td>Orthographic input lexicon</td>
<td>Orthographic input lexicon</td>
<td>Orthographic input lexicon</td>
</tr>
</tbody>
</table>

---

*PS is described in Hillis (1993); JJ is described in Hillis and Caramazza (1991); HG is described in Hillis (1991).† HG also had a semantic impairment for certain categories of words; this table describes her performance for categories for which she had intact comprehension, such as numbers and clothing. PS also had a mild semantic impairment in the categories of animals and vegetables; this table describes his performance for categories for which he had intact comprehension.
Table 2. Characteristics of Reported Individuals with Deep Alexia (Compensatory Use of GPC Mechanisms)

<table>
<thead>
<tr>
<th>Error types (example of errors)</th>
<th>KE*</th>
<th>RGB*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written word comprehension</td>
<td>Impaired (e.g., shoe understood as “mitten”)</td>
<td>Intact (e.g., six read as “seven,” but defined as “half a dozen”)</td>
</tr>
<tr>
<td>Spoken word comprehension</td>
<td>Impaired (e.g., “shoe” understood as mitten)</td>
<td>Impaired</td>
</tr>
<tr>
<td>Oral naming</td>
<td>Impaired (e.g., shoe named as “mitten”)</td>
<td>Impaired</td>
</tr>
<tr>
<td>GPC mechanisms</td>
<td>Impaired</td>
<td>Phonological output lexicon and GPC mechanisms</td>
</tr>
<tr>
<td>Levels of deficit</td>
<td>Semantic system and GPC mechanisms</td>
<td></td>
</tr>
</tbody>
</table>

*KE is described in Hillis et al. (1990); RGB is described in Caramazza and Hillis (1990).

...are misread, and the accuracy of performance on related tasks, such as word comprehension and naming. The various components of the reading system have distinct neural substrates, so that damage to different parts of the brain results in different patterns of alexia (see Black and Behrmann, 1994, and Hillis et al., 2002, for reviews).

—Argye E. Hillis

References


Further Readings


Alzheimer’s Disease

Alzheimer’s disease (AD) is a neurodegenerative condition that results in insidiously progressive cognitive decline. According to widely recognized clinical diagnostic criteria for AD (McKhann et al., 1984), these patients have language processing impairments as well as difficulty with memory, visual perceptual-spatial processing, and executive functioning. The language impairment changes as the disease progresses (Bayles et al., 2000), and one profound consequence of language difficulty in AD is that this deficit strongly reflects clinical decline and the need for additional skilled nursing support (Chan et al., 1995). This article briefly summarizes the studies showing that AD patients’ language deficit includes difficulty with comprehension and expression of the sounds/letters, words, and sentences that are used to communicate in day-to-day circumstances. Work relating these language deficits to a specific neuroanatomical distribution of disease is also reviewed.

We may consider first the semantic impairment in AD. This deficit limits the comprehension and expression of concepts represented by single words and sentences. In expression, for example, a significant word-finding deficit is a prominent and early clinical feature of AD (Bayles, Tomoeda, and Trosset, 1990; White-Devine et al., 1996; Cappa et al., 1998). This is seen in spontaneous speech as well as on measures of confrontation naming. Naming difficulty due to a semantic impairment often is manifested as semantic paraphasic errors, such as the substitution of “chair” for the intended target “table.” As the condition progresses, AD patients’ spontaneous speech becomes limited to the use of overlearned phrases and ultimately becomes quite empty of content, while they fail to provide any responses during confrontation naming.

Semantic deficits are also prominent in comprehension. More than 50% of AD patients differ significantly from healthy seniors in their performance on simple category judgment tasks. For example, many AD patients are impaired when shown a word or a picture and asked, “Is this a vegetable?” (Grossman et al., 1996). Priming is a relatively automatic measure of semantic integrity: AD patients have relatively preserved priming for high-frequency lexical associates such as “cottage-cheese” that have little semantic relationship (Nebes, 1989; Ober et al., 1991), but they are impaired in their priming for coordinates taken from the same semantic category, such as “peach-banana” (Glosser and Friedman, 1991; Glosser et al., 1998). Item-by-item analyses show reduced priming for words that are difficult to understand and name (Chertkow, Bub, and Seidenberg, 1989). The unity of impairment across comprehension and expression is emphasized by the observation of the greatest naming difficulty in patients with significant semantic comprehension deficits (Chertkow and Bub, 1990; Hodges, Salmon, and Butters, 1992). The basis for this pattern of impaired semantic memory has been an active focus of investigation. Some studies associate the semantic comprehension impairment in AD with the degradation of knowledge about a word and its associated concept (Gonnerman et al., 1997; Garrard et al., 1998; Conley, Burgess, and Glosser, 2001). A category-specific deficit understanding or naming natural kinds such as “animals” compared with manufactured artifacts such as “implements” may emerge in AD (Siliveri et al., 1991). Other recent work suggests that difficulty understanding words and pictures in AD is related to an impairment in the categorization process that is so crucial to understanding concepts. In particular, AD patients appear to have difficulty implementing rule-based processes for understanding the critical features of words that determine category membership (Grossman, Smith, et al., submitted) or for learning the category membership of new concepts (Koenig et al., 2001).

The comprehension and expression of concepts often requires appreciating the long-distance relationships among several words in a sentence. Some early work attributes sentence processing difficulty in AD to a grammatical deficit. Paragrammatic errors such as “mices” and “catched” can be observed in speech, oral reading, and writing. Other studies relate impaired comprehension to difficulty with the grammatical features of phrases and a deficit in understanding grammatically complex sentences such as those containing a center-embedded clause (Emery and Breslau, 1989; Kontiola et al., 1990; Grober and Bang, 1995; Croft, Hodges, and Patterson, 1999). Essentially normal performance during on-line studies of sentence comprehension cast doubt on this claim (Kempler et al., 1998; Grossman and Rhee, 2001). More recently, considerable evidence indicates that sentence processing difficulty is related to a limitation in the working memory resources often needed to support sentence processing. Although the precise nature of the limited cognitive resource(s) remains to be established, AD patients’ grammatical comprehension deficit can be brought out by experimental manipulations that stress cognitive resources such as working memory, inhibitory control, and information processing speed. Studies demonstrate working memory limitations through the use of verbs featuring unusual syntactic-thematic mapping in a sentence comprehension task (Grossman and White-Devine, 1998),
concurrent performance of a secondary task during sentence comprehension (Waters, Caplan, and Rochon, 1995; Waters, Rochon, and Caplan, 1998), and limited inhibition of the context-inappropriate meaning of a polysemous word (Faust et al., 1997).

Semantic memory and sentence processing appear to be preserved in some AD patients. Nevertheless, this cohort of AD patients may have a different language impairment profile. Many of these patients, despite preserved single-word and sentence comprehension, are impaired in retrieving words from the mental lexicon. This kind of naming difficulty is marked by changes in the sounds contributing to a word, such as omissions and substitutions (Biassou et al., 1995). This limitation in lexical retrieval appears to be equally evident on oral lexical retrieval tasks and in writing. AD patients appear to be quite accurate at discriminating between speech sounds that vary in the place of articulation or voice onset timing, but phonemic (single-sound) substitutions can also be heard in their speech.

Alzheimer’s disease is a focal neurodegenerative condition. Functional neuroimaging studies obtained at rest with modalities such as single-photon emission computed tomography, positron emission tomography, and functional magnetic resonance imaging (fMRI) (Foster et al., 1983; Friedland, Brun, and Budson, 1985; DeKosky et al., 1990; Johnson et al., 1993; Alsop, Detre, and Grossman, 2000) and histopathological studies of autopsied brains (Brun and Gustafson, 1976; Arnold et al., 1991; Braak and Braak, 1995) show that specific brain regions are compromised in AD. The neuroanatomical distribution of disease revealed by these studies includes gross defects such as atrophy and microscopic abnormalities such as neuritic plaques and neurofibrillary tangles in the temporal, parietal, and frontal association cortices of the brain. The neural basis for the language difficulties in AD is investigated most commonly through brain-behavior correlation studies, although occasional functional neuroimaging reports describe defects in regional brain activation during language challenges. Early correlation studies relate sentence comprehension difficulty to reduced resting activity in posterior temporal and inferior parietal regions of the left hemisphere (Haxby et al., 1985; Grady et al., 1988). More recent work associates difficulty understanding single words and impaired confrontation naming with left temporoparietal cortex (Desgranges et al., 1998). Moreover, the defect in this brain region is significantly greater in AD patients with a semantic memory impairment than in AD patients with relatively preserved semantic memory (Grossman et al., 1997), and a comparative study demonstrates the specificity of this correlative pattern in AD relative to patients with a frontotemporal form of dementia (Grossman et al., 1998). By comparison, only very modest correlations show a relationship between grammatical comprehension and left inferior frontal cortex in AD.

A handful of functional neuroimaging studies report monitoring the regional cortical responses of AD patients during language challenges. One study shows limited activation of middle and inferior frontal regions in AD patients that had been recruited during a category membership semantic decision in healthy seniors (Saykin et al., 1999). More recently, a BOLD fMRI study of semantic judgments described limited activation of left temporoparietal cortex and frontal cortex in AD patients compared to healthy seniors, and AD patients recruited brain regions adjacent to the activated areas seen in elderly control subjects for specific categories of knowledge such as “animals” and “implements” (Grossman, Koenig, et al., in press).

AD patients thus have prominent deficits at several levels of language processing. This includes impaired semantic memory, manifested in measures of comprehension and expression. There is also difficulty with lexical retrieval in reading and writing, although perceptual judgments of speech sounds are relatively preserved. The neural basis for these language impairments appears to be a defect in temporoparietal association cortex of the left hemisphere; although a defect in left frontal association cortex also may contribute to the language impairments in AD.

See also DEMENTIA.

Acknowledgments

Work was supported in part by U.S. Public Health Service grants AG15116, AG17586, and NS35867.

—Murray Grossman

References


Global aphasia is an acquired language disorder characterized by severe loss of comprehension with concomitant deficits in expressive abilities (Peach, 2001). Unlike other syndromes of aphasia, few distinctions are found between preserved and impaired components of these patients’ language. The outlook for recovery from global aphasia tends to be bleak (Kertesz and McCabe, 1977). For this reason, the term global may be more prognostic than descriptive.

Several isolated areas of relatively preserved comprehension have been identified in global aphasia. These include recognition of specific word categories (Wapner and Gardner, 1979) and famous personal and geographic names (Yasuda and Ono, 1998). Globally aphasic subjects also show relatively better comprehension for personally relevant information (Van Lancker and Nicklay, 1992).

Globally aphasic patients are most severely impaired in their expressive abilities. The verbal output of many patients with global aphasia consists primarily of stereotypic recurring utterances or speech automatisms. Recurrent utterances have been described as being either nondictionary verbal forms (unrecognizable) consisting of consonant-vowel (CV) syllables (for example, do-do-do or ma-ma-ma) or dictionary forms (word or sentence) (Alajouanine, 1956). Blanken, Wallesch, and Papagno (1990) investigated the relationship between the nondictionary forms of recurrent utterances and comprehension disturbances in global aphasia. Although recurrent utterances are frequently associated with comprehension disturbances, the overall variability in language comprehension suggests that speech stereotypes cannot be used to infer the presence of severe comprehension deficits.

Patients with global aphasia give the impression of having more preserved communicative abilities than is actually the case because of their use of the suprasegmental aspects of speech. To investigate this, deBlesser and Poecsk (1985) analyzed the spontaneous utterances of a group of globally aphasic subjects with output limited to CV recurrences. They concluded that the length and pitch of these utterances were stereotypical and that the prosody of these patients did not seem to reflect communicative intent. The contributions to conversation that are credited to these patients, therefore, may be the result of the communicative partners’ need for informative communication rather than the patients’ use of prosodic elements to convey intent.

The efficiency of communication following global aphasia depends on the type of question that is asked (Herrmann et al., 1989). Better performance is observed for responses to yes/no questions than for responses to interrogative pronoun questions and narrative requests. Patients with global aphasia mostly use gesture in their responses to yes/no questions. The other types of questioning require increased verbal output and thus create the need for more complex communicative responses.

Patients with global aphasia rarely take the initiative to communicate or expand on shared topics (Herrmann et al., 1989). Their most frequent communication strategies are those that enable them to secure comprehension (e.g., indicating comprehension problems, requesting support for establishing comprehension). Although these individuals rely most heavily on nonverbal strategies, the efficiency of their communication may approximate that of less impaired aphasic patients while imposing nearly as low a burden on the communication partner (Marshall, Freed, and Phillips, 1997).

Global aphasia results most commonly from a cerebrovascular event in the middle cerebral artery at a level inferior to the point of branching. The majority of lesions producing global aphasia are extensive and involve both prerolandic and postrolandic areas of the left hemisphere. These include Broca’s (posterior frontal) and Wernicke’s (superior temporal) areas and may extend to subcortical areas, including the basal ganglia, internal capsule, and thalamus (Murdoch et al., 1986). Occasionally, the lesion is confined to anterior, posterior, or deep cortical and subcortical regions (Mazzocchi and Vignolo, 1979). Global aphasia has also been described in patients with lesions restricted to subcortical regions, including the basal ganglia, internal capsule, periventricular white matter, temporal isthmus, and thalamus (Alexander, Naeser, and Palumbo, 1987).

Ferro (1992) investigated the influence of lesion site on recovery from global aphasia. The lesions in his subjects with global aphasia were grouped into five types with differing outcomes. Type 1 included patients with large pre- and postrolandic middle cerebral artery infarcts. These patients had a very poor prognosis. The remaining four groups were classified as follows: type 2, prerolandic; type 3, subcortical; type 4, parietal, and type 5, double frontal and parietal lesion. Patients in these latter groups demonstrated variable outcomes, improving generally to Broca’s or transcortical aphasia. Complete recovery was observed in some patients with type 2 and 3 infarcts. In contrast, Basso and Farabola (1997) investigated recovery in three cases of aphasia based on the patients’ lesion patterns. One patient had global aphasia from a large lesion involving both the anterior and posterior language areas, while two other patients had Broca’s and Wernicke’s aphasia from lesions restricted to either the anterior or posterior...
language areas, respectively. The patient with global aphasia recovered better than his two aphasic counterparts and had an outstanding outcome. Basso and Farrabola concluded that group recovery patterns based on aphasia severity and site of lesion may not be able to account for the improvement that is occasionally observed in individual patients.

Global aphasia has the lowest recovery rate of all the aphasias (Kertesz and McCabe, 1977). When assessing the language recovery that does occur, comprehension is found to improve more than expression (Prins, Snow, and Wagenaar, 1978). Differences have been reported in the temporal patterns of recovery depending on whether the subjects were receiving speech and language treatment. For globally aphasic patients not receiving treatment, improvement appears to be greatest in the first 6 months after onset (Pashek and Holland, 1988).

Globally aphasic patients receiving treatment demonstrate substantial improvements during the first 3–6 months but also continue to improve during the period between 6 and 12 months or more after onset. Sarno and Levita (1981) observed the most accelerated improvement between 6 and 12 months after stroke. Nicholas et al. (1993) found different patterns of recovery for language and nonlanguage skills during the first year. Substantial improvements in praxis and oral-gestural expression were noted only in the first 6 months after onset, while similar improvements in auditory and reading comprehension were observed only between 6 and 12 months after onset.

The majority of patients with global aphasia will not recover to less severe forms of the disorder. Some patients, however, will improve such that the condition evolves into other aphasia syndromes, including Broca’s, transcortical motor, mixed nonfluent, conduction, anomic, and Wernicke’s aphasias. Occasionally, patients make a complete recovery to normal language. One apparent explanation for the variability among these patients might be the greater instability of language scores (and therefore aphasia classifications) obtained during the first 4 weeks after stroke versus those obtained after the first month post onset. McDermott, Horner, and DeLong (1996) found greater magnitudes of change in scores and frequencies of aphasia type evolution in subjects tested during the first 30 days after onset than in subjects tested in the second 30 days after onset. Aphasia tends to be more severe during the acute stage, giving observers an initial impression of global aphasia. However, globally aphasic patients who do progress to some other form of aphasia may demonstrate changes that extend into the first months after onset (Pashek and Holland, 1988). In some cases, the global aphasia may not begin to evolve until after the first month has passed. The discrepancies in these studies, therefore, do not appear to be simply the result of the time at which the initial language observations were recorded. Apparently, evolution from global aphasia is the result of a complex interaction among a number of heretofore incompletely understood variables.

Several factors have been investigated for their prognostic significance with regard to global aphasia. The patient’s age appears to have an impact on recovery: the younger the patient, the better the prognosis (Holland, Swindell, and Forbes, 1985). However, numerous exceptions to this trend have been described. Age may also relate to the type of aphasia at 1 year post stroke. In the study by Holland, Swindell, and Forbes (1985), younger globally aphasic patients evolved to a nonfluent Broca’s aphasia while older patients evolved to increasingly severe fluent aphasias with advancing age. The oldest patients remained globally aphasic.

Absence of hemiparesis with global aphasia may be a positive indicator for recovery. Tranel et al. (1987) described globally aphasic patients with dual discrete lesions (anterior and posterior cerebral) that spared the primary motor area. Global aphasia improved significantly in this group within the first 10 months after onset. These conclusions are tempered by the results of Keyserlingk et al. (1997), who found that chronic globally aphasic patients with no history of hemiparesis did not fare any better with regard to language outcome than did their globally aphasic counterparts with hemiparesis from the time of onset.

The radiologic findings of patients with global aphasia have also been studied to determine whether lesion patterns found on computed tomography may provide prognostic information. Although the findings have been generally mixed, Naeser et al. (1990) were able to show significantly better recovery of auditory comprehension for a group of globally aphasic subjects whose damage did not include Wernicke’s area (i.e., the lesions were limited to the subcortical temporal isthmus).

Finally, it appears that a lack of variability between auditory comprehension scores and other language scores may be viewed as a negative indicator (Mark, Thomas, and Berndt, 1992). The more performance differs among language tasks, the better the outlook. Within auditory comprehension scores, globally aphasic patients who produce yes/no responses to simple questions, regardless of their accuracy, seem to have a better outcome at 1 year post onset than those who cannot grasp the yes/no format.

—Richard K. Peach

References


Further Readings


Aphasia, Primary Progressive

The clinical syndrome of primary progressive aphasia is a diagnostic category applied to conditions in which individuals exhibit at least a 2-year history of progressive language deterioration not accompanied by other cognitive symptoms and not attributable to any vascular, neoplastic, metabolic, or infectious diseases. The disease
Primary progressive aphasia (PPA) is characterized by the gradual worsening of language dysfunction in the context of preserved memory, judgment, insight, visuospatial skills, and overall comportment, at least until the later stages of the disease. Historically viewed as an atypical presentation of dementia, the isolated deterioration of language in the context of degenerative disease has been reported for more than 100 years. PPA was first recognized as a distinct clinical entity by Mesulam in 1982. Mesulam and Weintraub (1992) distinguished PPA from other degenerative neurological conditions such as Alzheimer’s disease (AD) by its gradual progression of language dysfunction in the absence of more widespread cognitive or behavioral disturbances for a period of at least 2 years. In some cases, the syndrome may later become associated with cognitive and behavioral disturbances similar to those seen in frontal lobe or frontotemporal dementia or with motor speech disorders (dysarthria, apraxia of speech), as observed in corticobasal degeneration or upper motor neuron disease (e.g., primary lateral sclerosis) (Kertesz and Munoz, 1997). In other cases, the symptoms remain limited to the dissolution of language production and comprehension abilities throughout the duration of the affected individual’s life.

The diagnosis of PPA is typically made based on a 2-year history of progressive language deterioration that emerges in the absence of any marked disturbance of other cognitive function and is not associated with any vascular, neoplastic, metabolic, or infectious disease (Duffy and Petersen, 1992; Mesulam and Weintraub, 1992). In addition to neurological examination, medical assessment typically includes neuroimaging and neuropsychological testing. During the first few years, computed tomographic and magnetic resonance imaging results typically are negative or reveal mild to moderate atrophy of the left perisylvian region. Metabolic neuroimaging (e.g., positron emission tomography) typically reveals left perisylvian hypometabolism and is sensitive to abnormalities earlier than structural neuroimaging methods (e.g., Kempler et al., 1990; McDaniel, Wagner, and Greenspan, 1991). Neuropsychological assessment typically reveals relative preservation of nonverbal cognitive function (e.g., abstract reasoning, visual short-term memory, visuospatial organization) in conjunction with below-normal performance on tests requiring verbal processing, such as immediate verbal recall, novel verbal learning, and verbal fluency (e.g., Sapin, Anderson, and Pulaski, 1989; Weintraub, Rubin, and Mesulam, 1990). Additionally, many studies have reported the presence of nonlinguistic sequelae known to frequently co-occur in nonprogressive forms of aphasia, such as acalculia, dysphagia, depression, limb apraxia, and apraxia of speech (Rogers and Alarcon, 1999).

At present, the cause of PPA is unknown. It is similarly unclear whether PPA is related to a distinct neuropathological entity. Investigations of the neuropathology associated with PPA yield heterogeneous findings (Black, 1996). Most case studies that include histological data report nonspecific neuronal loss in the left perisylvian region accompanied by spongiform changes in the superficial cortical layers (Snowden et al., 1992; Scheltens, Ravid, and Kamphorst, 1994). Other cases have been reported of individuals who initially presented with progressive language disturbances but eventually were diagnosed with AD (Pogacic and Williams, 1984; Kempler et al., 1990), Pick’s disease (Holland et al., 1985; Graff-Radford et al., 1990; Scheltens et al., 1990; Kertesz et al., 1994), and Creutzfeldt-Jakob disease (Shuttleworth, Yates, and Paltan-Ortiz, 1985; Mandell, Alexander, and Carpenter, 1989). However, most of these cases do not meet the diagnostic criteria for PPA according to Mesulam’s definition. Furthermore, it is possible that the onset of neuropathophysiological changes related to AD develop sometime after the onset of the focal cortical degeneration associated with PPA, thus explaining the initial appearance of isolated language symptoms, followed by the onset of more widespread cognitive involvement. Thus, autopsy findings of pathology associated with AD do not preclude the possibility that two distinct disease processes may co-occur within the same individual. Although PPA is recognized as a distinct clinical entity, the issue of whether it is a distinct pathological entity remains unresolved.

The course of the disease is quite varied. After a 2-year history of isolated language symptoms, some proportion of individuals diagnosed with PPA eventually exhibit more widespread cognitive involvement consistent with a diagnosis of dementia (i.e., deterioration in two or more cognitive areas such as memory, personality changes, and the ability to independently carry out activities of daily living due to cognitive as opposed to physical impairments). Mesulam (1982) suggested waiting 5 years after the onset of symptoms before trying to predict the course of cognitive involvement. However, there is no indication from the literature that after 5 years, individuals are less likely to develop dementia (Rogers and Alarcon, 1999). Estimates vary concerning the percentage of individuals who, after initially presenting with a 2-year history of isolated language dissolution, eventually exhibit widespread cognitive involvement. These estimates range from 30% to 50% (Duffy and Petersen, 1992; Mesulam and Weintraub, 1992; Rogers and Alarcon, 1999). Thus it is likely that between 50% and 70% of individuals diagnosed with PPA experience only the consequences of declining speech, language, and communication for many years. These individuals continue to drive, manage their own finances, and in all respects other than speech, language, and communication maintain baseline levels of performance on repeated testing over many years. Thus, the course is variable, with some patients progressing rapidly, but for others, the course can be quite prolonged,
typically taking 6 or 7 years before they develop severe aphasia or mutism. For these individuals, the global cognitive deterioration does not occur as early in the disease process or to the same extent as seen in AD.

Researchers have attempted to identify clinical symptoms that could serve as reliable predictors of eventual cognitive status. The profile of speech and language dysfunction has been investigated as a prognostic indicator of whether an individual is likely to develop generalized dementia. The hypothesis that a language profile consistent with fluent aphasia as opposed to nonfluent aphasia predicts a course of earlier cognitive decline has been investigated through case study (e.g., Snowden et al., 1992) and systematic review of the literature (Duffy and Petersen, 1992; Mesulam and Weintraub, 1992; Rogers and Alarcon, 1999). The fluency dimension in aphasia refers to a dichotomous classification based on the nature of the spoken language disturbance. Individuals with fluent aphasia produce speech at normal to fast rates with normal to long phrase length and few, if any, phonological speech errors. Verbal output in fluent aphasia is characterized as logorrheic (running on and on) and neologistic (novel words), and it tends to be empty (devoid of meaningful content). Disturbances of auditory comprehension and anoma (i.e., word retrieval difficulties) were the primary symptoms in most cases of fluent PPA reviewed by Rogers and Alarcon (1999). Hodges and Patterson (1996) found these to be the primary presenting complaints in the individuals with fluent PPA that they labeled semantic dementia. Unlike individuals with AD, individuals with semantic dementia or fluent PPA have relatively spared episodic (both recent and old autobiographical) memory but exhibit loss of semantic memory, especially as concerns mapping concepts to their spoken form (Kertesz and Munoz, 1997).

A nonfluent profile is characterized by effortful speech, sparse output with decreased phrase length, impaired access to phonological word-form information, infrequent use of grammatical markers, and disturbed prosody, and is frequently associated with apraxia of speech. Auditory comprehension, while affected, generally deteriorates later than expressive language skills. Nonfluent spontaneous speech and the production of phonemic paraphasias in naming have been proposed as important characteristics distinguishing PPA from the aphasia-like symptoms in AD. However, the language disorder evinced by individuals with PPA rarely fits neatly and unambiguously into the fluency typology. Snowden et al. (1992) described a group of individuals with PPA who exhibited expressive and receptive disruptions of phonology and semantics. Be\'land and Ska (1992) described an individual with PPA who presented with “a syntactic deficit as in Broca’s . . . auditory comprehension deficits of the Wernicke’s aphasia type . . . and phonemic approximations as found in conduction aphasia” (p. 358). Although there is no accepted classification for individuals exhibiting this profile, the term “mixed aphasia” has been applied (e.g., Snowden et al., 1992). Although reliable sorting of aphasia in PPA into the fluent or nonfluent classes has not been established, the hypothesis that individuals with a fluent profile are more likely to develop generalized cognitive involvement than those with a nonfluent profile has received much attention (e.g., Weintraub, Rubin, and Mesulam, 1990; Duffy and Peterson, 1992; Snowden et al., 1992; Rogers and Alarcon, 1999).

The hypothesis that the profile of language impairment may predict the course of generalized cognitive involvement has been investigated primarily through systematic review of the literature. Duffy and Peterson (1992) reviewed 28 reports, published between 1977 and 1990, describing 54 individuals with PPA. Approximately half of the 54 individuals developed generalized dementia, but none of the 12 patients identified with nonfluent profiles evinced generalized cognitive involvement. This finding was interpreted as supporting the hypothesis that a nonfluent profile may predict a longer duration of isolated language symptoms, or perhaps a lower probability of developing widespread cognitive involvement. Mesulam and Weintraub (1992) reviewed 63 cases of PPA. The average duration of isolated language symptoms was 5.2 years, and six individuals exhibited isolated language symptoms for more than 10 years. They reported that, compared to either probable or pathologically confirmed AD, the PPA group contained more males, a higher incidence of onset before age 65, and a greater incidence of nonfluent aphasia. A nonfluent profile was never observed in the AD group, whereas the distribution of fluent and nonfluent profiles in the PPA group was balanced (48% fluent, 44% nonfluent). According to Mesulam and Weintraub (1992), not all individuals with probable AD exhibit aphasic disturbances, but those who do, exhibit only the fluent aphasia subtype.

Rogers and Alarcon (1999) reviewed 57 articles published between 1982 and 1998 describing 147 individuals with relatively isolated deterioration of speech and language for at least 2 years. Thirty-seven patients had fluent PPA, 88 had nonfluent PPA, and in 22 cases the type of aphasia was indeterminate. Among the individuals with fluent PPA, 27% exhibited dementia at the time of the published report. Among the nonfluent PPA group, 37% were reported to have developed generalized dementia. Of the 22 individuals with an undetermined type of aphasia, 73% exhibited clinical symptoms of dementia. The average duration of isolated language symptoms among the 77 individuals who developed generalized dementia was 5 years (6.6 years in fluent PPA, 4.3 years in nonfluent PPA, and 3.7 years among those with an undetermined type of aphasia). The aggregate data in this review did not support the hypothesis that individuals with a nonfluent profile are less likely to develop generalized cognitive involvement than those with a fluent profile. Furthermore, the data did not support the hypothesis that a nonfluent profile predicts a longer duration of isolated language symptoms. Despite the unequal number of patients in each of the fluency groups, the lack of control regarding the time post onset
across cases, and the possibility that there may be considerable impetus to report PPA in individuals who do not exhibit generalized dementia, it does not appear that the fluency profile is a reliable predictor of eventual cognitive status.

The initial symptoms of PPA vary from individual to individual, but anoma is the most commonly reported presenting complaint in patients with both fluent and nonfluent PPA (Mesulam, 1987; Rogers and Alarcon, 1999). Another early symptom, particularly in nonfluent PPA, is slow, hesitant speech, frequently punctuated by long pauses and filler words (“um,” “uh”). Although this may represent simply one of many manifestations of anoma, it also portends the language formulation difficulties that later render the speech of these individuals telegraphic (reduced mean length of utterance consisting primarily of content words). Impaired access to phonologic form is frequently associated with later-emerging spelling difficulties, although partial access to initial letters and syllable structure may be retained for many years (Rogers and Alarcon, 1998, 1999). Difficulties with phonologic encoding have also been reported in cases of fluent PPA. Tyler et al. (1997) described the anomic difficulties of one individual with fluent PPA as impaired mapping between the semantic lexicon and output phonology. More typically, individuals who eventually exhibit fluent PPA initially complain of difficulties understanding spoken language (Hodges and Patterson, 1996), whereas individuals with nonfluent PPA typically exhibit preserved language comprehension in the early stages (Karbe, Kertesz, and Polk, 1993). Some individuals with a nonfluent presentation of PPA initially exhibit motor symptoms consistent with a diagnosis of dysarthria or apraxia of speech. The initial symptom of progressive speech apraxia has been reported by Hart, Beach, and Taylor (1997). Dysarthria and orofacial apraxia have been reported as initial symptoms in a variation of PPA labeled slowly progressive anarthria (Broussolle et al., 1996). The relationships between and among nonfluent PPA, primary progressive apraxia, slowly progressive anarthria, corticobasal degeneration, primary lateral sclerosis, and Parkinson’s disease is of interest, because these conditions exhibit considerable clinical overlap and in some cases share similar pathophysiology. In the later stages of all of these syndromes, individuals lose the ability to communicate by speech and are uniformly described as “mute,” despite apparent differences regarding the underlying nature of the specific impairment precluding the production of spoken language.

Regardless of the subtype of PPA, the progressive loss of language need not result in the total cessation of all communication, as there are augmentative and alternative communication tools and strategies that can be proactively established so that the individual with PPA can maximize communication competency at every stage, despite the relentless deterioration in speech and language (Rogers, King, and Alarcon, 2000).

—Margaret A. Rogers

References


Aphasia: The Classical Syndromes

Aphasia is an acquired disorder of language subsequent to brain damage that affects auditory comprehension, reading, oral-expressive language, and writing. Early observations by Broca (1861a, 1861b) and Wernicke (1874) suggested that aphasia might be classified into a variety of syndromes, or types, based on differences in auditory comprehension and oral-expressive language behaviors. Moreover, different syndromes were believed to result from different sites of brain damage. Revisions of early classification systems yield a contemporary taxonomy that comprises seven syndromes: global, Broca’s, transcortical motor, Wernicke’s, transcortical sensory, conduction, and anomic (Benson, 1988; Kertesz, 1979).

Classification is based on the aphasic person’s auditory comprehension, oral-expressive fluency (phrase length and syntax), spoken repetition, and naming abilities. The seven syndromes can be divided into nonfluent, those with short phrase length and impaired morphosyntax (global, Broca’s, and transcortical motor), and fluent, those with longer phrase length and apparent preservation of syntactic structures (Wernicke’s, transcortical sensory, conduction, and anomic). An aphasic person’s syndrome may be determined by informal examination or by administering a standardized test, for example, the Western Aphasia Battery (WAB) (Kertesz, 1982) or the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass and Kaplan, 1983). The following describes each syndrome and the assumed site of lesion associated with each.

**Global Aphasia.** This nonfluent syndrome is associated with a large left hemisphere lesion that may involve the frontal, temporal, and parietal lobes, insula, and underlying white matter, including the arcuate fasciculus (Dronkers and Larsen, 2001). It is the most severe of all of the syndromes. Auditory comprehension is markedly reduced and may be limited to inconsistent comprehension of single words. Oral-expressive language is sparse, often limited to a recurring intelligible—“bees, bees, bees”—or unintelligible—“doobe, doobe, doobe”—stereotype. Other automatic expressions, including profanity and counting, may also be preserved. Globally aphasic patients are unable to repeat words, and no naming ability is present. Reading and writing abilities are essentially absent.

**Broca’s Aphasia.** This nonfluent syndrome receives its name from the early reports by Paul Broca (1861a, 1861b). Classical localization of the lesion resulting in Broca’s aphasia is damage in the left, inferior frontal gyrus—Broca’s area (Brodmann’s areas 44 and 45) (Damasio, 1992). However, both historical (Marie, 1906) and contemporary (Mohr, 1976; Dronkers et al., 1992) reports question the classical lesion localization. Patients have been described who have lesions in Broca’s area without Broca’s aphasia, and other patients have Broca’s aphasia but their lesion does not involve Broca’s area. Auditory comprehension is relatively good for single words and short sentences. However, comprehension of grammatically complex sentences is impaired. Their phrase length is short, and they produce halting, telegraphic, agrammatic speech that contains, primarily, content words. For example, describing how he spent the weekend, a patient with Broca’s aphasia related, “Ah, frat, no Saturday, ah, frisk, no, fishing, son.” Repetition of words and sentences is poor. Naming ability is disrupted, and reading and writing show a range of impairment.

**Transcortical Motor Aphasia.** Lichtheim (1885) provided an early description of this nonfluent syndrome, and he observed that the site of lesion spared the perisylvian language region. Currently, it is believed that the lesion resulting in transcortical motor aphasia is smaller than that causing Broca’s aphasia and is in the left anterior-superior frontal lobe (Alexander, Benson, and Stuss, 1989). With one exception, language behaviors are similar to those in Broca’s aphasia: good auditory comprehension for short, noncomplex sentences; short, halting, agrammatic phrase production; disrupted naming ability; and impaired reading and writing. The exception is relatively preserved ability to repeat phrases and sentences. Essentially, patients with transcortical motor
aphasia repeat much better than would be predicted from their disrupted, volitional productions.

Wernicke’s Aphasia. This fluent syndrome received its name from the early report by Carl Wernicke (1874). The traditional belief is that Wernicke’s aphasia results from a lesion in Wernicke’s area (posterior Brodmann’s area 22) in the left hemisphere auditory-association cortex (Damasio, 1992), with extension into Brodmann’s areas 37, 39, and 40. However, Basso et al. (1985) have reported cases of Wernicke’s aphasia resulting from exclusively anterior lesions, and Dronkers, Redfern, and Ludy (1995) have found Wernicke’s aphasia in patients whose lesions also spared Wernicke’s area. Spoken phrase length averages six or more words, and a semblance of syntax is present. However, the oral-expressive behavior includes phonological errors and jargon. One patient with Wernicke’s aphasia described where he went to college, Washington and Lee University, by relating, “There was the old one, ah Frulich, and the young one, young hunter, ah, Frulich and young hunter or Brulan.” A salient sign in Wernicke’s aphasia is impaired auditory comprehension. These patients understand little of what is said to them, and the deficit cannot be explained by reduced auditory acuity. In addition, verbal repetition and naming abilities are impaired, and there is a range of reading and writing deficits.

Transcortical Sensory Aphasia. This fluent syndrome may result from lesions surrounding Wernicke’s area, posteriorly or inferiorly (Damasio, 1992). Oral-expressive language is similar to that seen in Wernicke’s aphasia: longer phrase length and relatively good syntax. Auditory comprehension is impaired, similar to that in Wernicke’s aphasia, and naming, reading, and writing deficits are present. The salient sign in transcortical sensory aphasia is preserved verbal repetition ability for words and, frequently, long and complex sentences. Essentially, transcortical sensory aphasia patients repeat better than one would predict based on their impaired auditory comprehension.

Conduction Aphasia. Wernicke (1874) described this fluent syndrome. Lesion localization has been controversial. Geschwind (1965) proposed that conduction aphasia results from a lesion in the arcuate fasciculus that disrupts connections between the posterior language comprehension area and the anterior motor speech area. Damasio (1992) suggested that conduction aphasia results from damage in the left hemisphere supramarginal gyrus (Brodmann’s area 40), with or without extension to the white matter beneath the insula, or damage in the left primary auditory cortices (Brodmann’s areas 41 and 42), the insula, and the underlying white matter. Dronkers et al. (1998) reported that all of their patients with conduction aphasia had a lesion that involved the posterior-superior temporal gyrus, often extending into the inferior parietal lobule. The salient sign in conduction aphasia is impaired ability to repeat phrases and sentences in the presence of relatively good auditory comprehension and oral-expressive abilities. Although auditory comprehension is relatively good, it is not perfect. And, while oral-expressive language is fluent (longer phrase length and a semblance of syntax), patients with conduction aphasia make numerous phonological errors and replace intended words with words that sound similar. Naming, reading, and writing abilities are disrupted to some extent.

Anomic Aphasia. This fluent syndrome is the least severe. Anomia—word-finding difficulty—is present in all aphasic syndromes; thus, localization of the lesion that results in anomic aphasia is not precise. It can be found subsequent to anterior or posterior lesions (Dronkers and Larsen, 2001), and Kreisler et al. (2000) report anomic aphasia resulting from a lesion in the thalamus; medial temporal area; or frontal cortex, insula, and anterior part of the temporal gyri. Patients with anomic aphasia display longer phrase length and preserved syntax; mild, if any, auditory comprehension deficits; good repetition ability; and mild reading and writing impairment. Frequently, the anomic patient will substitute synonyms for the intended words or replace the desired word with a generalization, for example, “thing” or “stuff.”

Cautions

The classification of aphasia into the classical syndromes is not exempt from controversy. Some (Caramazza, 1984; Caplan, 1987) have challenged its validity. Darley (1982) suggested that aphasic people differ on the basis of severity or the presence of a coexisting communication disorder, frequently apraxia of speech. He advocated viewing aphasia unmodified by adjectives. The relationship between the site of lesion and the corresponding syndrome is also controversial. The classical sites of lesion for most aphasic syndromes are challenged by exceptions (Basso et al., 1985; Murdoch, 1988; Dronkers and Larsen, 2001). Some of the inconsistency may result from the time post onset when behavioral observations are made. Improvement in aphasia over time results in approximately 50% of aphasic patients changing from one syndrome to another (Kertesz and McCabe, 1977). Thus, an acutely aphasic patient with an inferior left frontal gyrus lesion may display the expected Broca’s aphasia; however, at 6 months after onset, the same patient’s language characteristics may resemble anomic aphasia. Confusion may also result from the methods employed to classify the aphasias. For example, classifications made with the WAB do not always agree with those made with the BDAE (Wertz, Deal, and Robinson, 1984). Finally, controversy and confusion may result from misuse of the term syndrome (Benson and Ardila, 1996). The behavioral profile that constitutes a specific aphasic “syndrome” is characterized by a range of impairment and not by identical performance among all individuals within a specific syndrome. In many, certainly not all, aphasic people, impaired behavioral features—fluency, auditory comprehension, verbal
repetition, naming—tend to result in different clusters that represent different profiles. These have led to the development and use of the classical syndromes in aphasia.

—Robert T. Wertz, Nina F. Dronkers, and Jennifer Ogar

References


Further Readings


A new concept in aphasiology was created when Wernicke (1874/1977) described ten patients with different forms of aphasia, and showed that two of the patients had fluent but paraphasic speech with poor comprehension (i.e., sensory aphasia). At autopsy of another patient, a lesion was found in the left posterior temporal lobe. This type of aphasia has been called by many names, including receptive, impressive, sensory, or more generally fluent aphasia. In most of the current classification systems, this type of syndrome is called Wernicke's aphasia. It affects 15%–25% of all patients with aphasia (Laska et al., 2001).

Although the exact boundaries of Wernicke's area are controversial, the typical lesion associated with Wernicke's aphasia is most often located in the posterior temporal area. The middle and superior temporal lobe posterior to the primary auditory cortex are affected in almost all cases. The primary auditory cortex is also often affected, as are the white matter subjacent to the posterior temporal lobe, the angular gyrus, and the supramarginal gyrus. In rare cases, restricted subcortical lesions may result in Wernicke's aphasia and hemiplegia, the latter being uncommon in cases with cortical lesions. Recent studies have not changed these classical views of the clinico-anatomical relations of initial aphasia.

Patients with Wernicke's aphasia are usually older than patients with Broca's aphasia. However, some rare cases of children with acquired fluent aphasia and a posterior temporal lesion have been described (Paquier and Van Dongen, 1991). Ferro and Madureira (1997) have attributed the age difference between patients with fluent aphasia and those with nonfluent aphasia to the higher prevalence of posterior infarcts in older patients. The most common etiological factor in vascular Wernicke's aphasia is cardiac embolus, which more often affects the temporal area, whereas carotid atherosclerotic infarctions are in most cases located in the frontoparietal area (Harrison and Marshall, 1987; Knepper et al., 1989). Coppens (1991), however, points to a higher mortality rate in older patients with stroke, which might cause a selection bias in studies showing a relationship between age and type of aphasia.

The typical clinical signs of Wernicke's aphasia include poor comprehension of spoken and written language and fluent but paraphasic (phonemic and semantic) speech. In some cases, neologistic jargon may occur. Naming is also severely affected, and phonemic or semantic prompting is of no help. Poor repetition distinguishes Wernicke’s aphasia from transcortical sensory aphasia. Writing mirrors the speech output. Handwriting is usually well formed, but the text is without content, and jargonagaphasia may occur. Because of posterior lesions, hemiparesis is present in rare cases, but visual field defects are more common. Many patients also show signs of anosognosia, especially during the acute stage of the illness. In most cases, the use of gestural communication or pantomime is affected as well.

Patients not traditionally classified as having aphasia may also show language disturbances resembling Wernicke's aphasia, such as patients with schizophrenia, dementia, or semantic dementia, a fluent form of primary progressive aphasia.

Some authors suggest that Wernicke's aphasia is not a uniform entity but includes many variants. Forms of neologistic, semantic, and phonemic jargon and pure word deafness may all be grouped under Wernicke's aphasia. Pure word deafness is a rare disorder characterized by severe difficulties in speech comprehension and repetition with preservation of other language functions, including the comprehension of nonverbal sounds and music (Kirshner, Webb, and Duncan, 1981). However, when Buchman et al. (1986) reviewed 34 published cases, they were unable to find any really pure cases—that is, cases without any other more generalized perceptual disorders that could be classified as acoustic agnosia or mild language disorders such as paraphasia, naming difficulties, and reading and writing disorders. Most of the patients with “pure” word deafness have had bilateral temporal lesions, but some patients with unilateral left hemisphere lesions have been described (Takahashi et al., 1992).

Personality factors may play a role in the clinical expression of aphasia. In some views, jargon aphasia is not solely a linguistic deficit. Rochford (1974) suggested that a pathological arousal mechanism and lack of control were crucial to jargon aphasia. Weinstein and Lyerly (1976) suggested that jargon aphasia could emanate from abnormal adaptation to the aphasic speech disorder. They found a significant difference in premorbid personality between patients with jargon aphasia and those without jargon aphasia. Most of their patients with jargon aphasia had a strong premorbid tendency to deny illness or openly expressed fear of illness, indicating the importance of anosognosic features in jargon aphasia.

Linguistically, patients with Wernicke’s aphasia speak with normal fluency and prosody without articulatory distortions. They often provide long and fluent answers (logorrhoea) to simple questions. In fact, patients with Wernicke’s aphasia produce an equal number of words as persons without aphasia in spontaneous speech. However, they show less lexical variety, a high proportion of repetitions, and empty speech (Bates et al., 2001). This may give an impression of grammatically correct speech, but the meaning of the utterances is lost because of a high proportion of paraphasias and neologisms (Lecours and Lhermitte, 1983). This type of speech error is called paragrammatism. Patients with Wernicke’s aphasia show morphological errors, but less so than...
patients with Broca’s aphasia (Bates et al., 2001). However, there is some evidence that in highly inflected languages such as Finnish, the number of errors is higher on inflected words than on the lexical stems (Niemi, Koivuselkä-Sallinen, and Laine, 1987). At least in spontaneous speech, distorted sentence structure in utterances of patients with Wernicke’s aphasia is related to the lexical-semantic difficulties rather than to morphosyntactic problems (Helasvu, Klippi, and Laakso, 2001). The same has been found in sentence comprehension. Patients with Wernicke’s aphasia performed correctly only on sentences that did not require semantic operations (Pinango and Zurif, 2001). According to these findings, the deficit in phonemic hearing does not explain the nature of comprehension problems in patients with Wernicke’s aphasia.

Most patients show skill in pragmatic abilities, such as using gaze direction and other nonverbal actions in conversation. Unawareness of one’s own speech errors usually occurs initially in Wernicke’s aphasia, but some degree of auditory self-monitoring develops after onset, and patients then begin to use various self-repair strategies to manage conversation (Laakso, 1997). In contrast to self-repair sequences in nonaphasic speakers, these sequences are very lengthy and often unsuccessful.

The initial severity of the aphasia is considered the most important single factor in predicting recovery from aphasia. Wernicke’s aphasia is usually tantamount to severe aphasia. In a study by Ross and Wertz (2001), of all patients with aphasia, those with Wernicke’s aphasia and global aphasia showed the most severe impairment in language functions and communication. These patients showed only limited recovery when measured at the impairment level by the Boston Diagnostic Aphasia Examination (BDAE) and at the disability level by CADL. In addition to initial severity of aphasia, supramarginal and angular gyrus involvements seem to relate to poor recovery in comparison with cases without extension to the posterior superior temporal gyrus (Kertesz, Lau, and Polk, 1993).

Patients who have recovered from Wernicke’s aphasia have shown a clear increase in activation in the right perisylvian area, suggesting a functional reorganization of the language with the help of the right hemisphere (Weiller et al., 1995). However, Karbe et al. (1998) reported that increased activity in the right hemisphere was present in patients with poor recovery and reflected the large lesions in the left hemisphere. Patients with good recovery showed increased activation in the left hemisphere surrounding the damaged area.

The classification of aphasia depends strongly on the methods used in the assessment. The major diagnostic tests, such as the BDAE, the Western Aphasia Battery (WAB), or the Aachener Aphasic Test (AAT), have slightly different criteria for classification. For example, whereas the WAB assigns all patients to some aphasia classification, up to 70% of patients examined with the BDAE might be designated as having unclassified aphasia. Another issue that confuses classification is the time after onset at which the evaluation is done. Depending on the sample studied, more than half of patients with aphasia will show evolution to another type of aphasia during the first year after the onset of illness (Ross and Wertz, 2001). Patients with initial Wernicke’s aphasia will usually evolve to have a conduction or transcortical type of aphasia, and may evolve further to have anomic aphasia (Pashek and Holland, 1988). On the other hand, the condition of elderly patients with initial global aphasia tends to evolve to Wernicke’s aphasia during the recovery period, and the condition of younger patients evolves to Broca’s aphasia. This could explain why only one-third of patients with fluent aphasia and lesions in Wernicke’s area have a persisting aphasia, and only slightly more than half of patients with chronic Wernicke’s aphasia have lesions in Wernicke’s area (Dronkers, 2000).

—Matti Lehtihalmes

References


Bryans (Eds.), Aphasiology (pp. 76–108). London: Baillière Tindall.

Further Readings


Aphasia Treatment: Computer-Aided Rehabilitation

The role of technology in treating clinical aphasiology has been evolving since studies first demonstrated the feasibility of using computers in the treatment of aphasic adults. This journey began with remote access to treatment in rural settings using large computer systems over the telephone. There followed the introduction and widespread use of personal computers and portable computers, with the subsequent development of complex software and multimedia programs. This changing course is not simply the result of technological progress but represents greater understanding by clinicians and researchers of the strengths and limitations of computer-aided treatment for aphasia and related disorders.

Four common types of treatment activities are appropriate for presentation on a computer: stimulation, drill and practice, simulations, and tutorials. Stimulation activities offer the participant numerous opportunities to respond quickly and usually correctly over a relatively long period of time for the purpose of maintaining and stabilizing the underlying processes or skills, rather than simply learning a new set of responses. It is easy to design computer programs that contain a large database of stimuli, and then to control variables (e.g., word length) as a function of the participant’s response accuracy. Drill and practice exercises teach specific information so that
the participant can function more independently. Stimuli are selected for a particular participant and goal, and therefore authoring or editing options are required to modify stimuli and target responses. A limited number of stimuli are presented and are replaced with new items when a criterion is reached. Because response accuracy is the focus of the task, the program should present an intervention or cues to help shape the participant’s response toward the target response. Drill and practice programs are convergent tasks. Simulations (“micro-worlds”) create a structured environment in which a problem is presented and possible solutions are offered. Simulations may be simple, such as presenting a series of text describing a problem, followed by a list of possible solutions. Complex programs more closely simulate real-life situations by using pictures and sound. Simulations provide the opportunity to design divergent treatment tasks that more fully recruit real-life problem-solving strategies than those addressed by more traditional, convergent computer tasks, for example, by including several alternative but equally correct solutions. Whether computer simulations can improve communicative behavior in real-life settings remains to be tested. Tutorials offer valuable information regarding communication and quality of life to the family, friends, and others who can influence the aphasic patient’s world for the better. Computer tutorials present information commonly found in patient information pamphlets but in an interactive, self-paced, format. Tutorials can incorporate features of an expert system, in which information is provided in response to a patient/family profile.

Computers can be incorporated into treatment in three fundamentally different ways. Computer-only treatment (COT) software is designed to allow patients, as part of clinician-provided treatment programs, to practice alone at the computer, without the simultaneous supervision or direct assistance from clinicians. The operation of COT programs should be familiar and intuitive for the patients, particularly those who cannot read lengthy or complex text. The program may alter in a limited way elements of the task in response to patient performance, such as reducing the number of stimuli or presenting predetermined cues in response to errors (e.g., Seron et al., 1980; Katz and Nagy, 1984). As all possible cues and therapeutic strategies that may be helpful to every patient cannot be anticipated, intervention is commonly simplistic, inflexible, or nonexistent. Consequently, COT programs are usually convergent tasks (e.g., drills) with simple, obvious goals and, if effective, increase treatment efficiency as supplementary tasks designed to reinforce or help generalize recently learned skills.

Computer-assisted treatment (CAT) software is presented on a computer as the patient and clinician work together on the program. The role for the computer is limited to supportive functions (e.g., presenting stimuli, storing responses, summarizing performance). The clinician retains responsibility for the most therapeutically critical components, particularly designing, administering, monitoring, and modifying the intervention in response to the patient’s particular needs. This relation between clinician and computer permits considerable flexibility, thus compensating for limitations inherent in the COT approach. In addition to treatment programs written specifically for use with clinicians (e.g., Loverso, Prescott, and Selinger, 1992; Van de Sandt-Koenderman, 1994), other software, such as COT word processing or a variety of video game programs, and even some web-based activities, can be used in this manner as long as clinicians provide patients with the additional information needed to perform the task.

Augmentative communication devices (ACDs) in aphasia treatment usually refer to small computers functioning as sophisticated “electronic pointing boards.” Unlike devices used by patients with severe dysarthria or other speech problems, patients with aphasia and other disorders affecting language cannot type the words they are unable to speak. ACDs designed for these individuals may incorporate digitized speech, pictures, animation, and a minimum of text. To facilitate both expression and comprehension, some devices are designed to permit both communication partners to exchange messages. Although ACDs vary in design and organization, some devices allow modification of the organization and semantic content in response to the particular needs and abilities of each patient. Researchers such as Aftonomos, Steele, and Wertz (1997) claim that for some patients with aphasia, treatment utilizing ACDs results in improved performance on standardized tests and in “natural language” (speaking, listening, etc.).

A speech-language pathologist educated in communication theory and sufficiently experienced in the clinic and in real life can create an infinite number of novel and relevant treatment activities and evaluate and modify these activities in response to unique and idiosyncratic patient behavior, even when those behaviors are unanticipated, for example, resulting from previously unacknowledged associations. In contrast, computer-provided treatment is based on a finite set of rules that are stated explicitly to evoke specific response that are (at best) likely to occur at particular points during a future treatment session, as in a game of chess. However, unlike chess, many elements of language, communication, and rehabilitation are not well delineated or universally recognized.

In describing four interrelated properties of computers and programming, Bolter (1984) helped aphasiologists better understand the relation between computers and treatment. (1) Computers deal with discrete (or digital) units of data, typically unambiguous numbers or other values, but many fundamental and recurrent aspects of communication are not clearly defined or understood. Whether during treatment or real-life, purposeful interactions, language and communication units are often incomplete, emanate (simultaneously and sequentially) from various modalities, and depend on context and past experiences. (2) Computers are conventional, that is, they apply predetermined rules to symbols
that have no effect on the rules. Regardless of the value of the symbols, the sophistication of the program (“complex branching algorithms”) or the outcome of the program, the rules never change (e.g., Katz and Nagy, 1984). In aphasia treatment, all the rules of treatment are not known and those that are may not be correct for all conditions, requiring clinicians to monitor and sometimes modify rules for each patient. Unlike clinician-provided treatment, computer-provided treatment does not modify the rules of the treatment it applies; therefore, computers do not respond adequately to the dynamics of patient performance. (3) Computers are finite. Their rules and symbols are defined within the program. Except in a limited way for artificial intelligence software, unanticipated responses do not result in the creation of new rules and symbols. Therapy demands a different approach. Not all therapeutically relevant behaviors have been identified, and those that have often vary in relevance among patients and situations. Treatment software that incorporates artificial intelligence (e.g., Guyard, Masson, and Quiniou, 1990) only roughly approximates this approach, usually by reducing the scope of the task. (4) Computers are isolated from real-world experience. Problems and their solutions exist within the boundaries of the program and frequently have little to do with reality. Problems are created with the intention that they can be solved by manipulating symbols in a predetermined, finite series of steps. This lack of “world knowledge” is perhaps the most significant obstacle to comprehensive computer-provided treatment, as it limits the ability of programs to present real-world problems with multiple options and practical, flexible solutions.

In an extensive review of the literature, Robinson (1990) reported that the efficacy of computer-aided treatment for aphasia and for other cognitive disorders had not been demonstrated. The research studies reviewed suffered from inappropriate experimental designs, insufficient statistical analyses, and other deficiencies. Robinson stated that some researchers obscured the basic question by asking what works with whom under what conditions (see Darley, 1972).

There is no substitute for carefully controlled, randomized studies, the documentation of which has become the scientific foundation of aphasiology. Research reported over the last 15 years has assessed the effect of particular computerized interventions (e.g., Cerrar, Ellis, and Dean, 1996) and incorporated increasingly sophisticated designs and greater numbers of subjects to assess the efficacy of computer-aided aphasia treatment, from simple A-B-A designs and comparisons of pre- and posttreatment testing (Mills, 1982) to large, randomly assigned single-subject studies (Loverso, Prescott, and Selinger, 1992) and group studies incorporating several conditions (Katz and Wertz, 1992, 1997).

The efficacy of computerized aphasia treatment is being addressed one study at a time.

See also SPEECH AND LANGUAGE DISORDERS IN CHILDREN: COMPUTER-BASED APPROACHES.

—Richard C. Katz

References


Further Readings


Aphasia Treatment: Pharmacological Approaches

For over a century, clinicians have sought to use pharmacological agents to remediate aphasia and/or to help compensate for it, but without much success (Small, 1994). However, in several limited areas, the use of drug treatment as an adjunct to traditional (behavioral) speech therapy has shown some promise. Furthermore, the future for pharmacological and other biological treatments is bright (Small, 2000).

In this brief article, we restrict our attention to the subacute and chronic phases of aphasia, rather than the treatments for acute neurological injury. Much of this work has focused on a class of neurotransmitters, the catecholamines, which occur throughout the brain. Two important catecholamines are dopamine, produced by the substantia nigra, and norepinephrine, produced by the locus coeruleus. Since the catecholamines do not cross the blood-brain barrier, typical therapy involves agents that increase catecholamine concentrations. Dextro-amphetamine is the most popular agent of this sort, acting nonspecifically to increase the concentrations of all the catecholamines at synaptic junctions. In the early studies, a single dose of dextro-amphetamine led to accelerated recovery in a beam-walking task in rats with unilateral motor cortex ablation (Feeney, Gonzalez, and Law, 1982). By contrast, a single dose of haloperidol, a dopamine antagonist, blocked the amphetamine effect. When given alone, haloperidol delays spontaneous recovery, whereas phenoxybenzamine, an α1-adrenergic antagonist, reproduces the deficits in animals that have recovered. Similar results have now been obtained in several species and in motor and visual systems (Feeney and Sutton, 1987; Feeney, 1997).

The role of antidepressant medications in stroke recovery, including selective serotonin reuptake inhibitors (SSRIs) and the less selective tricyclics, is not straightforward. Neither fluoxetine (an SSRI) nor direct administration of serotonin seems effective in improving motor function in a rat model (Boyeson, Harmon, and Jones, 1994), whereas the tricyclics have produced mixed effects (Boyeson and Harmon, 1993; Boyeson, Harmon, and Jones, 1994).

The role of the inhibitory transmitter γ-aminobutyric acid (GABA) has been investigated in several studies. Intracortical infusion of GABA exacerbates the hemiparesis produced by a small motor cortex lesion in rats (Schallert et al., 1992). The short-term administration of diazepam, a benzodiazepine and indirect GABA agonist, can permanently impede sensory cortical recovery. Furthermore, phenobarbital, which may have some GABA agonist effects, also impedes recovery (Hernandez and Holling, 1994).

A number of early studies were conducted with limited success and are summarized in two recent reviews (Small, 1994; Small, 2001). Modern studies of pharmacological treatment of aphasia have focused on neurotransmitter systems, particularly catecholaminergic...
systems. A number of studies have been conducted, not all well designed. At present, no drug has been adequately shown to help aphasia recovery to the degree that would be necessary to recommend its general use. Several biological approaches that have been tested for aphasia recovery have been shown to be ineffective (e.g., meprobamate, hyperbaric oxygen) or very poorly supported by published results (e.g., amobarbital, selegiline).

Several studies have examined the role of dopamine. Albert et al. (1988) described a case suggesting that the dopamine agonist bromocriptine helped restore speech fluency in a patient with transcortical motor aphasia resulting from stroke. Another case report failed to find a similar benefit in a similar patient (MacLennan et al., 1991). Two additional patients improved in speech fluency but not in other aspects of language function (Gupta and Mlcoch, 1992). Another open-label study suggested some effect in moderate but not severe aphasia (Sabe, Leiguarda, and Starkstein, 1992).

Dextro-amphetamine is perhaps the most widely studied biological treatment for the chronic effects of stroke, including aphasia (Walker-Batson, 2000), yet both its clinical efficacy and mode of action remain unclear (Goldstein, 2000). Nonetheless, evidence from both animal model systems and humans make this a somewhat promising drug for the treatment of aphasia.

In a study of motor rehabilitation from stroke, more than half of a group of 88 elderly patients who had been classified as rehabilitation failures because of poor progress in physical therapy benefited from dextro-amphetamine as an adjunct to physical therapy (Clark and Mankikar, 1979). A double-blind placebo-controlled study replicated this finding (Crisostomo et al., 1988) in eight patients with ischemic stroke.

An early study of aphasia pharmacotherapy with methylphenidate (similar to amphetamine) and chlordiazepoxide (a benzodiazepine) revealed no effects (Darley, Keith, and Sasanuma, 1977). A recent prospective double-blind study of motor recovery with methylphenidate found a significant difference in motor and depression scores on some measures but not others (Grade et al., 1998). Methylphenidate may play a role in the treatment of post-stroke depression (Lazarus et al., 1992).

Walker-Batson et al. (1991) have reported a study of six aphasic patients with ischemic cerebral infarction. Each patient took dextro-amphetamine every 4 days, about an hour prior to a session of speech and language therapy, for a total of ten sessions. When evaluated after this period, the patients performed at significantly above expected levels.

Of potential significance, the studies showing beneficial effects of dextro-amphetamine, that is, the study by Walker-Batson et al. (1991), a motor study by the same group (Walker-Batson et al., 1995), and the other study of motor rehabilitation (Crisostomo et al., 1988), share the common feature of evaluating the drug as an enhancement to behavioral or physical therapy rather than as a monotherapeutic panacea.

Piracetam is a GABA derivative that acts as a nootropic agent on the central nervous system (CNS) and facilitates cholinergic and excitatory amine neurotransmission (Giurgea, Greindl, and Preat, 1983; Vernon and Sorkin, 1991). A large multicenter trial (De Deyn et al., 1997) showed no effect on the primary outcome measure of neurological status at 4 weeks. Another study showed improvement at 12 weeks that was no longer present at 24 weeks (Enderby et al., 1994). A later study (Huber et al., 1997) showed that improvement occurred on only one subtest (written language) of a large battery.

A crucial issue that must be addressed as part of aphasia rehabilitation is depression, since it can adversely affect language recovery. Following stroke, patients with depression have more cognitive impairment than patients with comparable lesions but no depression (Downhill and Robinson, 1994). Furthermore, in stroke patients matched for severity and lesion localization, patients with depression experience a poorer recovery than their nondepressed counterparts in functional status and cognitive performance (Morris, Raphael, and Robinson, 1992).

Growth factors have been advocated for a variety of purposes in the treatment of stroke, particularly in the acute phase of ischemic brain injury (Zhang et al., 1999), but also as neuroprotective agents useful in the chronic phase of recovery from brain injury (Olson et al., 1994). Gene transfer into the CNS might ultimately play a role in delivering trophins or other agents to damaged brain areas and thus to help stimulate recovery or increased synaptic connectivity.

Neural stem cells are multipotential precursors to neurons and glia. Attempts have been made to induce differentiation into neurons and glial cells, and further into specific types of such cells. Specifically with regard to stroke and the treatment of cortical lesions, fetal neocortical cells have been successfully transplanted into the site of cortical lesions (Johansson, 2000), and have even been shown to migrate selectively into areas of experimental cell death (Macklis, 1993; Snyder et al., 1997).

One important consequence of this research into the pharmacology of aphasia is the realization that drugs are not only potential therapeutic adjuncts but can also serve as inhibitors of successful recovery. The first study of this type, by Porch and colleagues (1985), showed that patients taking certain medicines performed more poorly on an aphasia battery than those who were not taking medicines.

In a formal retrospective (chart review) study, patients with motor deficits after stroke were divided into one group taking a number of specific drugs at the time of stroke (clonidine, prazosin, any dopamine receptor antagonist [e.g., neuroleptics], benzodiazepines, phenytoin, or phenobarbital) and another group that was not (Goldstein, 1995). Statistical analysis revealed that whereas patient demographics and stroke severity were similar between groups, motor recovery time was significantly shorter in the patients who were not taking one of these drugs.
This work has profound relevance to aphasia rehabilitation. To maximize functional recovery, it is important not only to ensure adequate behavioral treatment, but also to ensure the appropriate neurobiological substrate for this treatment (or, more concretely, to ensure that this substrate is not pharmacologically inhibited from responding to the therapy). It is thus advisable for patients in aphasia therapy to avoid drugs that might interfere with catecholaminergic or GABAergic function or that are thought to delay recovery by empirical study.

Current knowledge suggests a potential beneficial effect of increased CNS catecholamines on human motor recovery and aphasia rehabilitation. Although pharmacotherapy cannot be used as a replacement for speech and language therapy, it might play a role as an adjunct, and other biological therapies, such as cell transplantation, might play a role in concert with carefully designed, adaptive learning approaches. In the published cases where pharmacotherapy improved language functioning in people with aphasia, it was used adjunctively, not alone. It is very likely that pharmacotherapy has a valuable role to play as an adjunct to behavioral rehabilitation to decrease performance variability and to improve mean performance in patients with mild to moderate language dysfunction from cerebral infarctions.

Acknowledgments
Research was supported by the National Institute of Deafness and Other Communication Disorders, National Institutes of Health, under grant NIH DC R01-3378.
—Steven L. Small

References
Johansson, B. B. (2000). Brain plasticity and stroke rehabilita-
nervous system: Clinical experiences with NGF in Parkinson’s and Alzheimer’s diseases. *Journal of Neurology*, 242 (Suppl. 1), S12–S15.


**Aphasia Treatment: Psychosocial Issues**

Although researchers are developing a useful understanding of aphasia as a neurocognitive condition, the world we experience is a social and interactive one. The social aspects of life contribute to quality of life, although quality of life has been difficult to characterize scientifically and in ways that have clinical utility (Hilari et al., in press).

*Psychosocial* refers to the social context of emotional experience. Most emotions are closely associated with social interactions. Aphasia has implications for the individual’s whole social network, especially the immediate family. The value dimensions in our lives, such as health, sexuality, career, creativity, marriage, intelligence, money, and family relations, contribute to quality of life and are markedly affected for the aphasic person (Hinckley, 1989) and that person’s relatives. All may expect considerable disruption of professional, social, and family life, reduced social contact, depression, loneliness, frustration, and aggression (Herrmann and Wallesch, 1989).

Recovery and response to rehabilitation in aphasia are also significantly influenced by emotional and psychosocial factors. The aphasic person’s family and other caregivers need to be involved as much as possible in intervention, and this involvement extends beyond discharge. The experienced disability rather than the impairment itself is the focus of rehabilitation. Rehabilitation increasingly includes community-based work and support from not-for-profit organizations and self-help groups.

Whereas intervention during the acute stages of aphasia is largely based on the medical model, adjustment to aphasia is set more broadly within a social approach. Several broad psychosocial and quality-of-life areas have been incorporated into rehabilitation: dealing with depression and other emotions, social reintegration, and the development of autonomy and self-worth. Autonomy involves cooperating with others to achieve ends, whereas independence implies acting alone and may not be an achievable goal. These areas provide a basis for developing broad-ranging programs.

**Emotion**

We need to distinguish the *direct* effects of damage on the neurophysical substrate of emotion and the *indirect* effects, which are natural reactions to catastrophic personal circumstances (Code, Hemsley, and Herrmann, 1999). Our understanding of these factors is improving. Three different forms of depression can follow damage: catastrophic reaction, major post-stroke depression, and minor post-stroke depression. There is little research separating reactive from direct effects but Herrmann, Bartels, and Wallesch (1993) found significantly higher ratings for physical signs of depression, generally considered direct effects, during the acute stage.

Some view depression accompanying aphasia within the framework of the grief model (Tanner and Gertenberger, 1988). In this view, individuals grieve for the loss of communication, moving through the stages of denial, anger, bargaining, depression, and acceptance. Whether people do in fact go through these stages has not been investigated, but it has served as a framework for counseling. Determining denial, bargaining, acceptance, and so on is problematic but has been investigated interpretively with personal construct therapy techniques by Brumfit (1985), who argues that aphasia affects a person’s core role constructs, with grief for the essential element of *self* as a speaker.

A further problem is that the symptoms of depression, such as changes in sleep and eating, restlessness, and
crying, can also be caused by physical illness, anxiety, hospitalization, and factors unrelated to mood state. Language impairment plays a special role in the problem of identifying and measuring mood, and one approach has been to use the nonverbal Visual Analogue Mood Scale (VAMS), substituting schematic faces for words (Stern, 1999).

Aström et al. (1993) found major depression in 25% of stroke survivors, which rose to 31% at 3 months post onset, fell to 16% at 12 months, and increased again over the next 2 years to 29%. Some workers take the position that drugs should be avoided, but others suggest incorporating drugs and psychotherapy, with drugs perhaps being more appropriate at early stages to counter direct effects and psychotherapy and counseling more appropriate later, when the individual is more ready to deal with the future. Individual and group counseling are effective, and aphasic people themselves have recently become involved as counselors.

Social Reintegration and Self-Esteem

Self-esteem and self-worth are complex constructs, tied to social activity, that workers suggest should be central to psychosocial rehabilitation. The importance of self has been examined by Brumfit (1993) using personal construct techniques. Facilitating participation in the community entails passing responsibility to the individual gradually so that the individual can develop autonomy, develop greater self-esteem, and take greater ownership of the issues that they face. The importance of involving the aphasic individual fully has been addressed by Parr et al. (1997).

Hersh (1998) argues that particularly at discharge from formal therapy, an account of the ongoing management of psychosocial adjustment is needed. Simmons-Mackie (1998) argues that the traditional concept of a plateau being reached, where linguistic progress slows down for many, is not relevant when considering the social consequences of aphasia. The aim should be to prepare and assist clients to integrate into a community. Social affiliation is stressed as a means of maintaining and developing self-identity.

Not-for-profit organizations and centers such as the Pat Orato Aphasia Center in Ontario and the Aphasia Center in Oakland, California, provide community-based programs for aphasic people and their families, including the training of relatives and professionals as better conversation partners. Through such programs, the psychosocial well-being of both aphasic participants and their families is improved (Hoen, Thelander, and Worsley, 1997). Training volunteers as communication partners results in gains in psychological well-being and communication among aphasic participants, caregivers, and the communication partners themselves. In philosophy and approach, these centers resemble United Kingdom charities such as Speakability and Connect. Organizations like these are increasingly offering more long-term and psychosocially oriented programs. Many use volunteers, who figure increasingly in social reintegration, providing a valuable resource and helping to establish, facilitate, and maintain groups.

The efficacy of self-help groups in which the aphasic members decide on the group’s purpose, take responsibility for running the group, and serve as officers of the group (e.g., chair, secretary) is being evaluated, particularly in relation to the development of independence and autonomy. Structured support groups are of benefit. The self-help groups in the United Kingdom attract mainly younger and less severely impaired individuals and use little in the way of statutory resources (Code et al., 2001).

Wahrborg et al. (1997) have reported benefits from integrating aphasic people into educational programs and organizations, and Elman (1998) has introduced adult education instructors into an aphasia center.

Work and other purposeful activity is an important value dimension central in the development and maintenance of self-worth and autonomy. Returning to work remains a constant concern of many aphasic people. Ramsing, Blomstrand, and Sullivan (1991) explored prognostic factors for return to work, but there has been little follow-up to this research. Parr et al. (1997) report that only one person in their study who was working at the time of the stroke returned to the same employment. A few found part-time work, and the rest became unemployed or retired. Garcia, Barrette, and Laroch (2000) studied perceived barriers to work and found that therapists focused on personal and social barriers, employers focused on organizational ones, and aphasic people focused on barriers of all types. The groups also suggested strategies for reducing barriers to work.

Family therapy, to include people close to the aphasic person, is generally beneficial and can lead to positive changes (Nichols, Varchevker, and Pring, 1996).

Conclusions

A psychosocial approach to improving quality of life aims to aid social reintegration in such a way that the individual is able to maintain identity, develop self-esteem and purpose, and become socially re affiliated. Significant others, professionals, and volunteers are involved. This approach offers a challenge to clinicians, as it extends their role and requires a more comprehensive approach to management. There remains a lack of evidence-based approaches to managing psychosocial adjustment, but it is clear that volunteers, organizations, charities, and community-based centers are contributing. The challenge facing clinical aphasiology is to evaluate the benefits of psychosocial support for aphasic people and its impact on quality of life.

Acknowledgments

The first author was a Fellow at the Hanse Institute of Advanced Study, Delmenhorst, Germany, when this chapter was completed.

—Chris Code and Dave J. Muller
References


Further Readings


Aphasic Syndromes: Connectionist Models

Connectionist models of aphasic syndromes first emerged in the late nineteenth century. Broca (1861) described a patient, Leborgne, whose speech was limited to the monosyllable *tan* but whose ability to understand spoken language and nonverbal cues and ability to express himself through gestures and facial expressions were normal. Leborgne’s brain contained a lesion whose center was in the posterior portion of the inferior frontal convolution of the left hemisphere, now known as Broca’s area. Broca claimed that Leborgne had lost “the faculty of articulate speech” and that this brain region was the neural site of the mechanism involved in
speech production. In 1874, Karl Wernicke described a patient with a speech disturbance that was very different from that seen in Leborgne. Wernicke's patient was fluent, but his speech contained words with sound errors, other errors of word forms, and words that were semantically inappropriate. Unlike Leborgne, Wernicke's patient did not understand spoken language. Wernicke related the two impairments—one of speech production and one of comprehension—by arguing that the patient had sustained damage to the “storehouse of auditory word forms,” leading to speech containing the types of errors that were seen and impaired comprehension. By extrapolation from a similar case he had not personally examined, Wernicke concluded that the patient’s lesion was in the posterior portion of the left superior temporal gyrus, now known as Wernicke’s area, and that this region was the locus of the “storehouse of auditory word forms.” Wernicke argued that, in speaking, word sounds were conveyed from Wernicke’s area to Broca’s area, where the motor programs for speech were developed. This connection gave this type of model its name.

Lichtheim (1885) developed a more general model of this type. Lichtheim recognized seven syndromes, listed in Table 1. Lichtheim argued that these syndromes followed lesions in the regions of the brain depicted in Figure 1. These syndromes were critiqued on neuro-anatomical grounds (Marie, 1906; Moutier, 1908), dismissed as simplifications of reality that were of help only to schoolboys (Head, 1926), and ignored in favor of different approaches to language (Jackson, 1878; Goldstein, 1948). Nonetheless, they endured. Benson and Geschwind (1971) reviewed the major approaches to aphasia as they saw them and concluded that all researchers recognized the same basic patterns of aphasic impairments, despite using different nomenclature. Three more syndromes have been added by theorists such as Benson (1979), and Lichtheim’s model has been rounded out with specific hypotheses about the neuro-anatomical bases for several functions that he could only guess at.

Additional neuroanatomical foundation was first suggested in a very influential paper by Geschwind (1965). Geschwind argued that the inferior parietal lobe was a tertiary association cortical area that received projections from the association cortex immediately adjacent to the primary visual, auditory, and somesthetic cortices in the occipital, temporal, and parietal lobes. Because of these anatomical connections, the inferior parietal lobe served as a cross-modal association region, associating word sounds with the sensory qualities of objects. This underlay word meaning, in Geschwind’s view. Damasio and Tranel (1993) extended this model to

Table 1. Aphasic Syndromes Described by Lichtheim (1885)

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Clinical Manifestations</th>
<th>Hypothetical Deficit</th>
<th>Classical Lesion Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broca’s aphasia</td>
<td>Major disturbance in speech production with sparse, halting speech, often misarticulated, frequently missing function words and bound morphemes</td>
<td>Disturbances in the speech planning and production mechanisms</td>
<td>Posterior aspects of the 3rd frontal convolution (Broca’s area)</td>
</tr>
<tr>
<td>Wernicke’s aphasia</td>
<td>Major disturbance in auditory comprehension; fluent speech with disturbances of the sounds and structures of words (phonemic, morphological, and semantic paraphasias); poor repetition and naming</td>
<td>Disturbances in the permanent representations of the sound structures of words</td>
<td>Posterior half of the first temporal gyrus and possibly adjacent cortex (Wernicke’s area)</td>
</tr>
<tr>
<td>Pure motor speech disorder</td>
<td>Disturbance of articulation; apraxia of speech, dysarthria, anarthria, aphemia</td>
<td>Disturbance of articulatory mechanisms</td>
<td>Outflow tracts from motor cortex</td>
</tr>
<tr>
<td>Pure word deafness</td>
<td>Disturbance of spoken word comprehension, repetition often impaired</td>
<td>Failure to access spoken words</td>
<td>Input tracts from auditory system to Wernicke’s area</td>
</tr>
<tr>
<td>Transcortical motor aphasia</td>
<td>Disturbance of spontaneous speech similar to Broca’s aphasia with relatively preserved repetition; comprehension relatively preserved</td>
<td>Disconnection between conceptual representations of words and sentences and the motor speech production system</td>
<td>White matter tracts deep to Broca’s area connecting it to parietal lobe</td>
</tr>
<tr>
<td>Transcortical sensory aphasia</td>
<td>Disturbance in single-word comprehension with relatively intact repetition</td>
<td>Disturbance in activation of word meanings despite normal recognition of auditorily presented words</td>
<td>White matter tracts connecting parietal lobe to temporal lobe or portions of inferior parietal lobe</td>
</tr>
<tr>
<td>Conduction aphasia</td>
<td>Disturbance of repetition and spontaneous speech (phonemic paraphasias)</td>
<td>Disconnection between the sound patterns of words and the speech production mechanism</td>
<td>Lesion in the arcuate fasciculus and/or cortico-cortical connections between Wernicke’s and Broca’s areas</td>
</tr>
</tbody>
</table>
actions, arguing that associations between word sounds and memories of actions were created in the association cortex in the inferior frontal lobe. Geschwind (1965) and Damasio and Damasio (1980) also argued that the anatomical link between Wernicke's area and Broca's area (in which a lesion caused conduction aphasia) was the white matter tract known as the arcuate fasciculus.

The three and a half decades that have passed since publication of Geschwind's paper have brought new evidence for these syndromes and their relationships to brain lesions. Aphasic syndromes have been related to the brain using a series of neuroimaging techniques, first radionuclide scintigraphy with technetium, then computed tomography, magnetic resonance imaging, and positron emission tomography. All have confirmed the relationship of the major syndromes to lesion locations. These aphasic syndromes and their relationships to the brain figure prominently in recent reviews of aphasia in leading medical journals (e.g., Damasio, 1992).

Despite this revival, the connectionist approach to aphasic syndromes is under renewed attack.

A major limitation of the classical syndromes is that they stay at arm's length from the linguistic details of language impairments. The classical aphasic syndromes basically reflect the relative ability of patients to perform entire language tasks (speaking, comprehension, etc.), not the integrity of specific operations within the language processing system. Linguistic descriptions in these syndromes are incomplete and unsystematic. For instance, the speech production problem seen in Broca's aphasia can consist of one or more of a large number of impairments: dysprosodic speech, poorly articulated speech, agrammatism, an unusual number of short phrases. If all we know about a patient is that she or he has Broca's aphasia, we cannot tell which of these problems (or other) that person has.

A second problem is that identical deficits occur in different syndromes. For instance, certain types of naming problems can occur in any aphasic syndrome (Benson, 1979). Because of this, most applications of the clinical taxonomy result in widespread disagreements as to a patient's classification (Holland, Fromm, and Swindell, 1986) and to a large number of "mixed" or "unclassifiable" cases (Lecours, Lhermitte, and Bryans, 1983). The criteria for inclusion in a syndrome are often somewhat arbitrary: How bad does a patient's comprehension have to be for the patient to be identified as having Wernicke's aphasia instead of conduction aphasia, or global aphasia instead of Broca's aphasia? There have been many efforts to answer this question (see, e.g., Goodglass and Kaplan, 1972, 1982; Kertesz, 1979), but none is satisfactory.

A third problem with the classical aphasic syndromes is that they are not as well correlated with lesion sites as the theory claims they should be. These syndromes are related to lesion sites reasonably well only in cases of rapidly developing lesions, such as stroke. Even in these types of lesions, the syndromes are never applied to acute and subacute phases of the illness. Even in the chronic phase of diseases such as stroke, at least 15% of patients have lesions that are not predictable from their syndromes (Basso et al., 1985), and some researchers think this figure is much higher—as much as 40% or more, depending on what counts as an exception to the rule (de Bleser, 1988). We now know that the relationship between lesion location and syndrome is more complex than we had thought, even in cases in which the classical localization captures part of the picture. Broca's aphasia, for instance, does not usually occur in the chronic state after lesions restricted to Broca's area but requires much larger lesions (Mohr et al., 1978). Some theorists have argued that the localizing value of the classical syndromes reflects the co-occurrence of variable combinations of language processing deficits with motor impairments that affect the fluency of speech (Caplan, 1987; McNeil and Kent, 1991). From this point of view, the localizing value of the classical syndromes is due to the invariant location of the motor system.

Finally, the classical syndromes offer very limited help to the clinician planning therapy, because the syndromes give insufficient information about what is wrong with a patient. For example, knowing that a patient has Broca's aphasia does not tell the therapist what aspects of speech need remediation—articulation of sound segments, prosody, production of grammatical elements, formulation of syntactic structures, and so on. Nor does it guarantee that the patient does not need therapy for a comprehension problem; it only implies
that any comprehension problem is mild relative to the problems of other aphasics or to the patient’s speech problem. Finally, it does not guarantee that the patient does not have other problems, such as anosmia, difficulty reading, and the like. In practice, most clinicians do not believe that they have adequately described a patient’s language problems when they have identified that patient as having one of the classic aphasic syndromes. Rather, they specify the nature of the disturbance found in the patient within each language-related task; for example, they indicate that a patient with Broca’s aphasia is agrammatic, has a mild anomia, and so on. Detailed psycholinguistic and linguistic descriptions of aphasic impairments are slowly replacing the disconnection approach to syndromes.

It is a feature of the history of science, and some think a tenet of the philosophy of science, that people do not abandon a theory because it has inadequacies. Some philosophers of science think that no theory is ever proven wrong. According to this view, theories are abandoned because people get tired of them, and people get tired of theories because they have others that they think are better. This perspective on science applies to the connectionist approach to aphasic syndromes. The classic syndromes have not been abandoned, but their acceptance is waning, and there are new developments that address some of their inadequacies.

Acknowledgment

Work was supported in part by a grant from NIDCD (DC00942).

—David Caplan

References


Aphasiology, Comparative

Comparative aphasiology is the systematic comparison of aphasia symptoms across languages, including the comparison of acquired reading problems across languages and writing systems. The goal of study is to ensure that theories that claim to account for the various constellations of aphasic symptoms can handle the similarities and differences seen in aphasic speakers of languages of different types (see Menn, 2001, for a discussion of language typology in the context of comparative aphasiology). Serious experimental and clinical comparative work began in the 1980s; however, researchers long before that time understood that comparative data are essential. Otherwise, general theories of aphasia would depend on data from the few, closely related languages of the countries in Europe and North America where research in neurolinguistics was being undertaken: English, French, German, and, for a time, Russian.

The clearest example of such a premature class of theories was the ‘‘least pronunciation effort’’ approach
to agrammatism. This approach focused on the prevalence of bare-stem forms in agrammatic output in English, French, and German and suggested that such forms were used to avoid extra articulatory effort. A more phonologically sophisticated theory (Kean, 1979), relying on the same small database, proposed that agrammatic speakers were constrained to produce only the minimum phonological word. However, the languages then available for study have bare-stem forms that are not only short but also grammatically unmarked (mostly first- or third-person singular present tense of verbs and the singular of nouns) and very frequent. Therefore, the data could also support a markedness-based or a frequency-based account. Or the underlying problem could be morphological or morphosyntactic instead of articulatory or phonological.

A reason to think that the problem is actually morphological (problems with retrieving inflections) or morphosyntactic (problems with computing which inflections the syntax demands) is that endings of participles and infinitive forms are better preserved than other verb forms in English, Italian, French, and German. Is this because they have no tense marking or no person marking? Or is there another reason?

To test the hypothesis that a problem with inflections is a specific type of problem with morphosyntax, we look for languages where inflections are controlled in different ways. For example, if problems with verbs are blamed on difficulty with person or number agreement, we should see whether verb problems also exist in languages without agreement, such as Japanese. Or the problem may be deeper than morphosyntax: the verb problem could be due to a semantic difference between nominal and verbal types of elements, as suggested by experimental work on Chinese, which has no agreement or inflection (Bates, Chen, et al., 1991; Chen and Bates, 1998). Or perhaps multiple factors interact—a more difficult claim to test.

Another type of problem that demands a comparative approach is the question of why there are so few adjectives in aphasic language. From, say, a German-centered point of view, we might ask, Is this a conceptual problem, an agreement problem (nonexistent in English, so that could not be the sole problem), or a problem in inserting elements between article and noun (in which case, Romance languages, where most NPs have article-noun-adjective order, should not show the effect)?

Issue after issue requiring a comparative approach can be listed in the same way. Would relative clauses that do not require movement (as in Chinese or Japanese) be deployed using different mechanisms? (Jaeger et al., 1996; Pinker, 1999).

Recent History

By 1980, comparative language acquisition studies (e.g., the work published in Slobin, 1985–1995) were well established and provided both intellectual and logical models for comparative aphasiology. Bellugi’s team at the Salk Institute began to compare aphasic syndromes in hearing/speaking and deaf/signing individuals (Bellugi, Poizner, and Klima, 1989), and an international group coordinated by Bates at the University of California–San Diego began using psycholinguistic techniques to do cross-linguistic studies of English and Italian, eventually expanding to Chinese, Russian, Spanish, and other languages. At the Aphasia Research Center of the Boston University School of Medicine, a team focused on morphosyntax in agrammatic narratives (Cross Language Aphasia Study I, Menn and Obler, 1988, 1990a) created a standard elicitation protocol and began to collect data from speakers with agrammatic aphasia and matched controls. Data were collected on 14 languages: the non-Indo-European languages Mandarin Chinese, Finnish, Hebrew, and Japanese and the Indo-European languages Dutch, English, French, German, Hindi, Icelandic, Polish, Serbo-Croatian, and Swedish. Many of these languages (plus Hungarian and Turkish) are also represented in special issues of Brain and Language (1991, 41:2) and Aphasiology (1996, 10:6). Michel Paradis has led international work on bilingual aphasia, including the development of the extensive set of Bilingual Aphasia Tests; Paradis (2001), which was also published as a special issue of Journal of Neurolinguistics (14:2–4), includes contributions on non-Indo-European Basque and Hungarian and Indo-European Afrikaans, Catalan, Czech, Farsi (Persian), Friulian, Greek, and Spanish, as well as material on African-American English and more data on Finnish, Polish, Hebrew, and Swedish.

Methods of Study

Comparative studies raise special methodological issues because of the need to ensure that all materials pose comparable levels of difficulty across languages and cultures. Drawings acceptable in one country may be anomalous in another; words that are (apparently) translation equivalents may not be comparably frequent, and so on (Menn et al., 1996). Several chapters in Paradis (2001) point out the importance of allowing for the effects of bilingualism and multiple dialect use, which are present in most of the world’s population. The presentation of comparative production data requires an elaborated interlinear translation format, so that any reader, familiar with the language or not, can see what the subject actually said, what he or she should have said, and what the errors were. A widely used version is derived from the interlinear morphemic translation style codified
by Lehmann (1982). The example below is taken from Farsi data (Nilipour, 2000).

1. sar\[-am\]
2. man [be] dast [va] sar xombáre [xor-d]
3. I [to] hand [and] head\[-my\] shrapnel [hit:PAST:3SG]
4. I hand, head shrapnel [hit].

Line 2 is the original as spoken, but edited to remove hesitations and phonetic or phonemic errors. Omitted words are supplied in square brackets if they can be reconstructed from the context. Line 1 gives the target forms for each substitution error, placed on the first line directly above the incorrect form. The third line is a morph-by-morph translation into the language of publication, with standard abbreviations indicating case, tense, aspect, and so on, with abbreviations for affixes. Line 4 contains a “smooth” translation into the language of the publication. Lines 1 and 4 are as similar as possible in their degree and types of errors—i.e., equally agrammatic, equally paraphasic. Line 5, not used in all publications, identifies part of speech, and codes for functor (all lowercase) versus content word (first letter uppercase), making it easier to count members of these categories. This is done because so much psycholinguistic theorizing hinges on form class and on the content/functor distinction.

Comprehension Studies and Recent Findings

Bates’s group (e.g., Bates, Friederici, and Wulfeck, 1987a) has used a task in which hearers are presented with a string of two nouns and a transitive verb (e.g., “The cow the pencils kick”) and must decide which of two nouns (or noun phrases) is the agent. The string is presented in all orders, whether grammatical or not in the language in question, and often with conflicting cues from word order, animacy, and number agreement. Their key finding has been that people with aphasia show the same language-specific preferences for interpreting these strings—the same tendency to place more reliance on word order, animacy, or agreement in making their interpretations—as do speakers without neurological impairment. Thus, agrammatic aphasia, in particular, cannot be an eradication of grammar. Further confirmation of this claim comes from Serbo-Croatian: Lukatela, Shankweiler, and Crain (1995) used a slightly more natural picture-choice comprehension task to show that people with agrammatic aphasia were able to distinguish between subject-gap sentences (“The lady is kissing the man who is holding an umbrella”) and object-gap sentences (“The lady is kissing the man that the umbrella is covering”), even when these sentences were constructed with the same word order, so that the hearers had to rely only on case markers and agreement markers (subject-verb agreement, modifier-noun agreement, agreement between pronouns and their referents, etc.). (Stimuli that convey this complex syntactic structure without varying the word order cannot be constructed in most European languages.)

While these results showed that agrammatic aphasics could use morphological information, Bates’s group has shown that morphological cues are also the ones most likely to be underutilized by speakers with all forms of aphasia, as well as by control subjects who are loaded with competing experimental task demands (Blackwell and Bates, 1995).

Note that not all comparative studies make direct contrasts across two or more languages. A study is also comparative if it selects a language to work in specifically because that language enables us to tease apart variables of interest. The elaborate morphology and free word order of Serbo-Croatian allowed Lukatela and colleagues to examine comprehension of morphology with word order held constant. Similarly, gender agreement in French was used by Jakubowicz and Goldblum (1995, p. 242) to construct an experiment contrasting the preservation of grammatical morphemes in nonfluent aphasia. They found that “local” (within noun phrase) markings were better preserved than ones requiring computation across major syntactic boundaries. Luzzatti and De Bleser (1996) reported a similar result for production.

Production Studies and Findings

A variety of production studies have supported the following general claims. (1) The greater the semantic importance of a morpheme, the more likely that it will be produced. For example, although grammatical morphemes are in general prone to errors, and free grammatical morphemes tend to be omitted by speakers with nonfluent aphasias, negation is almost never omitted (Menn and Obler, 1990b; see also Friederici, Weissenborn, and Kail, 1991). (2) The larger the paradigm of choices for a given form, the more likely that errors will be made (Bates, Wulfeck, and MacWhinney, 1991; Paradis, 2001b). Aphasic production errors in morphosyntax tend to be only one semantic feature away from the target (gender or number or case or tense); the direction of errors is probabilistic, but more frequent forms are more likely to be produced correctly (Dressler, 1991). However, an individual may have a preferred “default” form not shared by other aphasic speakers of the same language (Magnúsþóttir and Thráinsson, 1991). (3) In paradigms with multiple stem forms, errors tend to keep the same stem as the target (Mimouni and Jarema, 1997). Semantically appropriate case forms may be chosen even when the verb or preposition that would control that case is not produced. Most errors, especially those of nonfluent aphasics, are misselections from existing paradigms, but a few, notably some instances in Basque, involve the creation of nonexistent forms from existing morphemes (Laka and Erriondo Korostola, 2001) or the production of nonexistent stem forms (Swedish: Månsson and Ahlsén, 2001).
(4) Classifier languages show classifier errors (Tzeng et al., 1991), with nonfluent aphasic speakers tending to omit them or to fall back on a "neutral" classifier; speakers with fluent aphasia tend to commit more substitution errors, as is characteristic of their production patterns across morpheme categories and languages.

Of course, widening the database also broadens the questions: Why are utterance-final particles preserved in Japanese aphasia but not so well in Chinese (Paradis, 2001b)? Why does canonical word order make comprehension and elicited production easier across languages, even those that have "free" word order, and why is some noncanonical order occasionally used as a production default? Cross-linguistic psycholinguistic and computational-linguistic approaches seem to be the most promising avenues to further understanding of aphasia, and, more generally, to the question of how language is represented in the human brain.

Note. Given the state of the art and the extent of individual variation within aphasia syndromes, I use the terminology of various authors cited, without attempting to differentiate between the overlapping diagnostic categories of Broca’s aphasia and agrammatic aphasia (agrammatism); “non-fluent” aphasia includes both of these categories, plus several others from the traditional clinical categories. “Fluent” aphasia includes anomic aphasia (amnesia) and Wernicke’s aphasia. Textbook anomic aphasia has fluent articulation, disfluencies due only to word-finding difficulties, and no comprehension problems; textbook Wernicke’s aphasia has word-finding problems, comprehension problems, fluent articulation, and possible use of empty speech or nonsense “words” to compensate for impaired lexical retrieval.

—Lise Menn

References


Further Readings


Argument Structure: Representation and Processing

In a Principles and Parameters syntax framework, sentences are derived by two operations, merger and movement. Merger takes two categories as input (e.g., V and NP) and merges them into a single, higher-order category (e.g., VP). There are, however, constraints on the categories that can be merged successfully. Consider the following pairs:

1a. [NP The girl] sneezed
1b. *[NP The girl] sneezed [NP the boy]

2a. [NP The girl] defeated [NP the boy]
2b. *[NP The girl] defeated

3a. [NP The girl] gave [NP the prize] [PP to [NP the boy]]
3b. *[NP The girl] gave [NP the prize]

The (a) examples above are well-formed sentences; the (b) versions, containing the same verbs but different structures following the verbs, are ill-formed. Thus, not all verbs can fit into all sentence structures. How, then, does a theory of syntax account for these facts? Borrowing from logic, we can say that sentences are composed of a verb (i.e., predicate) and a set of arguments. A verb denotes an activity or event and an argument denotes a participant in the activity or event. So, in the grammatical (a) versions above, the sentences contain the appropriate number of arguments the verb entails; in the ungrammatical (b) versions there is either an extra argument (as in (1b)) or a required argument is missing (as in (2b) and (3b)), hence violating the argument structure of the verb.

Not all the phrases in a sentence function as arguments of a verb. Consider:
4. [The girl] defeated [the boy] on the beach/this afternoon/with great finesse.

The bracketed NPs are clearly participants in the event denoted by defeated, and thus are arguments of the verb. However, the italicized phrases do not represent participants in the event. Instead, they carry additional information (i.e., where the event took place, when, and the manner in which it took place). These expressions are considered to be adjuncts; that is, they are adjunctive or in addition to the information specified by the verb. Simplifying a bit, adjuncts are typically optional while arguments are often required.

**Thematic Roles.** The NPs that are participants/arguments of the verb in (4) play different semantic roles in relation to the verb defeated. A more comprehensive account of argument structure, then, needs to consider a description of these roles. For example, in (4) the NP the boy is the Agent of defeat, while the girl is the affected object and hence the Patient (or Theme). Some common thematic roles, then, are Agent, Experiencer, Theme/Participant, Goal, and Location.

A verb (that is, a lexical category) assigns its thematic roles to its arguments through theta marking. For example, the verb defeat is said to theta-mark the subject argument with the Agent role and the object argument with the Theme (Patient) role:

5. defeat V
   (Agent Theme)
   [NP The girl] defeated [NP the boy]  *The girl defeated

Thus, each lexical category (e.g., verb) has a set of argument structure features that must be satisfied in the sentence in which the word appears. If those features are not satisfied, the sentence will be ungrammatical.

Verbs can also have clauses as arguments; these too need to be theta-marked:

6. know V
   (Experiencer Proposition)
   [NP The coach] knew [CP that the girl defeated the boy]

The verb know assigns Experiencer to the subject NP argument and Proposition to the CP argument.

**Processing.** Canonical linking rules have been hypothesized to play a key role in the acquisition (van der Lely, 1994) and the processing of such verb-argument structures (see, e.g., McRae, Spivey-Knowlton, and Tanenhaus, 1998). This linking or mapping refers to the regular, most frequent relation found between thematic roles and syntactic functions (Pesetsky, 1995). For example, if an individual knows that a verb involves an Agent, Patient/Theme, and Goal, she can infer that those arguments can serve the role of subject, object, and oblique object, respectively. The verb donate, for example, requires three arguments—a subject NP, a direct object NP, and an indirect (oblique) object NP—as in:

7. [The girl/AGENT] donated [the present/THEME] to [the boy/GOAL].

Unlike (7), where the properties of the verb donate entail canonical linking, there are verbs with properties that entail noncanonical linking. Consider, for example, receive, which entails a reversal of the canonical assignment of Agent and Goal arguments:

8. [The boy/GOAL] received [the present/THEME] from [the girl/AGENT].

Sentences (7) and (8) reflect a well-known bias that suggests the “sender” seems more volitional and is a more plausible candidate for the Agent role than a “receiver” (Dowty, 1991). Importantly, there are no positional or configurational distinctions between the arguments to signal the difference in thematic order in the above examples; that is, the underlying syntax between the constructions appears to be the same. Thus, the distinction is based on the inherent properties of the verb.

Unlike the sentences above, where thematic roles can be directly assigned from inherent lexical information and linking relationships, sentences with so-called displaced arguments require indirect thematic role assignment. For example, consider the following noncanonical cleft-object sentence: “It was the boy who the girl kissed yesterday.” The direct object NP the boy has been displaced from its canonical, post-verb argument position in the sentence, leaving a gap. Such constructions are often referred to as filler-gaps. Verb-argument structure properties influence such constructions rather directly; for example, in the cleft-object case, a verb must license a direct object argument position in order to form a filler-gap dependency.

Given its syntactic and semantic importance, then, argument structure (in various forms) has played a privileged role in accounts of language processing. One of the earliest attempts to show that such lexically based information has repercussions for normal adult sentence processing was that of Fodor, Garrett, and Bever (1968). They inserted verbs differing in grammatical complexity (defined, in current terms, by the types of arguments each allowed) into matched sentences and found that off-line performance on those sentences decreased when verbs were more complex. Similar effects were later found on-line (e.g., Shapiro, Zurif, and Grimshaw, 1987).

What is important here is not just the fact that there are observed “effects” of argument structure, but what those effects suggest about the architecture of the sentence processing system. Briefly, most current accounts claim that when a verb (or any theta-assigning head of a phrase, including prepositions) is encountered in a sentence, its various argument structure configurations are momentarily activated (see, e.g., Pritchett, 1992; Mac-

References


Further Readings

In his important treatise on attention and effort, Kahneman (1973) specified some defining attributes of attention. Among his attributes, he suggested that attention is a limited capacity commodity (whether viewed as a single or multiple pool system). Attention is mobile and can be shifted either through mechanisms of orienting, enduring dispositions, or through the executive control system. The distributor of processing resources allocates attention according to a policy that (1) is biased toward novel stimuli, (2) has the ability to allocate attention to a particular domain or message, and (3) operates as a function of externally generated arousal levels. That attention is limited in capacity has been a central organizing principle for much of the research in attention and has given rise to the “dual task” paradigm, a widely used research method for investigating attention. The dual task is an experimental procedure whereby two tasks are performed concurrently and some aspect of each task is manipulated independently. The tasks are frequently manipulated by having the subject voluntarily allocate different percentages of attention or effort to each task (e.g., 50%/50%, 25%/75%, 75%/25%, 100%/0%, 0%/100%). If attention is shared between the two tasks, a trading of performance levels is expected and is typically expressed as a performance operating curve (POC). While the validity of the voluntary allocation part of the design has been challenged (Gopher, Brickner, and Navon, 1982), an alternative method is frequently used in which the inherent difficulty of the two tasks is manipulated parametrically. Again, a trading of performance levels is expected if the two tasks share a common pool or source of attention, and a POC is plotted and measured in order to test this hypothesis. The dual task paradigm, however, is not the only or even the most widely used approach to study attention. Without doubt, the Stroop test (Stroop, 1935) is the most widely researched attention task. In this task, unwanted intrusion of information is assessed through the rapid identification of colors or words of stimuli that are either congruent (e.g., written word “red” in red print) or incongruent (e.g., written word “red” in blue print). In this task, the subject is required to identify either the word or the color, and accuracy and response times are measured. Naming the color of a written word in the incongruent condition produces poorer accuracy and longer response times than in congruent conditions, indicating a competition for activation and inhibition of linguistic and nonlinguistic intentions and stimuli. N400 evoked potentials (Kutas and Hillyard, 1980; Bentin, 1987; Holcomb, 1988) and functional imaging (e.g., Just et al., 1996) are also common methods used to assess the role of attention in language processing.

Experimental paradigms are not the only source of evidence that attention is a construct worthy of study. Introspection also has provided a motivation for entertaining the notion of attention and its relationship to language. Indeed, most adults have had the experience of having read several pages of written material only to discover that nothing of what was read was remembered because the mind had wandered and focused on review-
ing yesterday’s particularly puzzling diagnostic or a previous argument with a colleague or the dean, or had focused on planning an upcoming holiday, course lecture, or treatment plan. Likewise, most have discovered the need to turn the radio off when encountering a difficult driving condition or courteously requesting the children to refrain from their backseat banter while formulating a response to the patrolman approaching the car following an apparent traffic violation. Although it is quite intuitive that attention to both internal and external stimuli plays an important role in many (perhaps all) language tasks, major syntheses and analyses of the general attention literature (e.g., Lang, Simons, and Balaban, 1997; Pashler, 1998) and the neuropsychological deficits of attention (e.g., Cohen, 1993; van Zomeren and Brouwer, 1994) have failed to address the role of attention in language processing. Indeed, not one of these major texts devoted to attention have addressed the role of attention in developmental disorders of language such as specific language impairment (SLI), or in acquired disorders affecting language processing, such as those of aphasia or traumatic brain injury. Only very recently has this subject received space in edited books on language and aphasia (e.g., Nadeau, Rothi, and Crosson, 2001) and only relatively recently have theoretical formulations (McNeil, Odell, and Tseng, 1991) of how attention might account for language impairments, and summary reviews (Murray, 2000; Crosson, 2001a; Fischler, 2001) offered explanations or hypotheses of how attention might interact with language impairments.

Language knowledge is characterized by the information that is represented in the brain along with the rules that govern it. Linguistic theory attempts to account for the structure of the information (rules and representations) that is stored in memory. Psycholinguistic theory attempts to account for conditions under which the rules and representations are stored or accessed and the various ways in which the different components are combined to produce or comprehend sounds, morphophonemes, words, phrases, sentences, and discourse. Informing and directing the linguistic system requires each individual language user to engage in a finely tuned interplay between internally generated intentions, linguistic knowledge, and a massive amount of sensory information that is continuously available, in addition to the selection (planning), programming, and execution of appropriate responses. This temporally demanding interplay creates an astonishing array of factors that have to be sorted and managed at all instances in time and on a continuous basis. It is the domain and role of attention and resource allocation to account for the gating (inhibition) and activation of endogenous intentions and exogenous stimuli involved in the formulation, comprehension, and production of language. Indeed, the role of attention in normal language processing has a long history, and evidence supports the conclusion that all levels of language processing require and compete for attentional resources with other language processes and with the processing of nonlinguistic information. For example, the attentional demands placed on language processing have been illustrated for lexical/semantic processing through priming paradigms (Neely, 1977), through evoked potentials (Kutas and Hillyard, 1980), and through dual-task studies (Arvedson and McNeil, 1987; Murray, 2000). Attentional demands for language have also been demonstrated in dual-task studies for syntactic processing by Blackwell and Bates (1995), for phonemic processing by Tseng, McNeil, and Milenkovic (1993), for auditory prosody processing by Slansky and McNeil (1997), and between language and nonlinguistic tasks by LaPointe and Erickson (1991).

Disorders of language are common and account for a sizable proportion of all communication disorders. Within the various classification systems for language disorders, it is widely recognized that there are multiple causes. Most systems acknowledge deficits at the representational level, including the rules used to govern these representations. Deficits at this level are often referred to as deficits of linguistic competence. A variety of performance factors are also recognized that can cause an otherwise competent or intact linguistic system to malfunction. Examples of performance deficits include disorders of linguistic-specific memory processes (Baddeley, 1993; Crosson, 2001b) and slowed perceptual or cognitive mechanisms (Tallal, Stark, and Mellits, 1985). Disorders of various aspects of the attentional system include orienting of attention (Robin and Rizzo, 1989), selective attention (Petry et al., 1994; Murray, Holland, and Beeson, 1998), inability to engage or disengage attention (Posner, Snyder, and Davidson, 1980), and resource allocation (McNeil, Odell, and Tseng, 1991). The construct of attentional deficits underlying language deficits is neither new (e.g., Kreindler and Fradis, 1968) nor restricted to aphasia. Campbell and McNeil (1985), for example, illustrated attentional deficits in an acquired pediatric language disorder population, and Barkley (1996) has applied the construct to attention deficit-hyperactivity disorder. However, restricting the discussion to the language impairment in aphasia, the past decade has seen a renewed interest in various aspects of attention, but primarily in the allocation of processing resources.

While skepticism remains apparent in some circles, it is widely recognized that explanations of language and other domains of cognition (e.g., memory, learning, executive function) that fail to account for attentional phenomena will remain incomplete. This is especially true for those areas of cognitive dysfunction resulting from brain damage (congenital or acquired, regardless of the time or cause of the injury) and developmental disabilities.

—Malcolm R. McNeil

References

with left and right hemisphere lesions. Clinical Aphasiology, 17, 188–201.


Further Readings


**Auditory-Motor Interaction in Speech and Language**

Carl Wernicke argued that cortical areas involved in the sensory representation of speech played an important role in speech production (Wernicke, 1874/1969). His argument was based on the clinical observation that the speech output of aphasic patients with posterior lesions in the left hemisphere was fluent but error prone. Modern evidence from both lesion and neuroimaging studies strongly supports Wernicke’s claim, and recent work has made progress in identifying an auditory-motor interface circuit for speech.

Developmental considerations make a strong case for the existence of an auditory-motor integration network for speech (Doupe and Kuhl, 1999). In learning to articulate the speech sounds in the local linguistic environment, there must be a mechanism by which (1) sensory representations of speech uttered by others can be stored, (2) articulatory attempts can be compared against these stored representations, and (3) the degree of mismatch revealed by this comparison can be used to shape future articulatory attempts. This auditory-motor integration network is still functional in adults, as revealed by the fact that it is possible to repeat pseudo-words accurately and by the effects of late-onset deafness on speech output (Waldstein, 1989).

Clinical evidence supports the view that “sensory” cortex participates in speech production. The classical fluent aphasias—Wernicke’s aphasia, conduction aphasia, transcortical sensory aphasia, and anomic aphasia—are all associated with left posterior cerebral lesions, that is, with regions that are commonly thought to be sensory in nature. Yet each of these fluent aphasias has prominent speech output symptoms: semantic and/or phonemic paraphasias (speech errors), paragrammatism (inappropriate use of grammatical markers), and anomia (naming difficulties) (Damasio, 1992). This observation demonstrates the general point that posterior “sensory” systems play an important role in speech production.

Evidence relevant to the more specific issue of auditory-motor integration comes from conduction aphasia (Hickok, 2001). A hallmark of conduction aphasia is the predominance of phonemic paraphasias, which can be evident across a large range of production tasks, including spontaneous speech, naming, reading aloud, and repetition (Goodglass, 1992). The preponderance of phonemic errors has led some authors to characterize conduction aphasia as a selective impairment in phonological encoding for production (Wilshire and McCarthy, 1996). Although the classical model holds that conduction aphasia is a disconnection syndrome involving damage to the arcuate fasciculus (Geschwind, 1965), recent evidence has shown that the syndrome can be caused by damage to, or electrical stimulation of, auditory-related cortical fields in the left superior temporal gyrus (Damasio and Damasio, 1980; Anderson et al., 1999). This region has been strongly implicated in speech perception, based on neuroimaging data (Zatorre et al., 1996; Norris and Wise, 2000), suggesting some degree of overlap in the systems supporting sensory and motor aspects of speech. This argument raises an apparent paradox, namely, that damage to systems strongly implicated in speech perception (i.e., left superior temporal gyrus) leads to a syndrome, conduction aphasia, characterized predominantly by a production deficit. This paradox can be resolved, however, on the assumption that speech perception is largely bilaterally organized, and that residual abilities of right hemisphere auditory systems function sufficiently well to support auditory comprehension (Hickok, 2000; Hickok and Poeppel, 2000).

Recent neuroimaging studies have supported and extended findings from the clinical literature. The left superior temporal gyrus has been shown to activate during a variety of speech production tasks (where speech is produced covertly, so that there is no external auditory input) including picture naming (Levelt et al., 1998; Hickok et al., 2000), repetition (Buchsbaum, Hickok, and Humphries, 2001), and word generation (Wise et al., 1991). Importantly, evidence from an MEG study of picture naming (Levelt et al., 1998) has shown that this left superior temporal activation occurs during a time frame prior to articulatory processes, suggesting
that this region is involved in phonological code retrieval in preparation for speaking and is not merely a form of motor-to-sensory feedback mechanism, although the latter mechanism may also exist.

Two studies have looked explicitly for overlap in activation associated with speech perception and speech production. The first used positron emission tomography to map areas of overlap when participants listened to stories versus performed a verb generation task (Papathanassiou et al., 2000). A region of overlap was found in the superior temporal gyrus, predominantly on the left, as expected, based on results reviewed earlier. Additional areas of overlap included inferior temporal regions and portions of the left inferior frontal gyrus. The second study (Buchsbaum et al., 2001) used functional magnetic resonance imaging to map activated regions when subjects first listened to and then covertly rehearsed a set of three multisyllabic pseudo-words. Two left posterior sites responded both to the auditory and motor phases of the trial: a site in the sylvian fissure at the parietal-temporal boundary (area Spt) and a more ventral site in the superior temporal sulcus. Brodmann’s area 44 (posterior Broca’s area) and a more dorsal premotor site also responded to both the auditory and motor phases of the trial. The activation time course of area Spt and of area 44 in that study were particularly strongly correlated, suggesting a tight functional relation. A viable hypothesis is that the STS site supports auditory representations of speech and that the Spt site serves as an interface system translating between auditory and motor representations of speech. This hypothesis is consistent with recent work in vision demonstrating the existence of visuomotor systems in the dorsal parietal lobe that compute coordinate transformations, such as transformations of retinocentric to head- and body-centered coordinates, which allows visual information to interface with various motor-effector systems that act on that visual input (Andersen, 1997; Rizzolatti, Fogassi, and Gallese, 1997).

Sensorimotor interaction is pervasive across many hierarchical levels in the central nervous system. The empirical record supports conceptual arguments for sensorimotor interaction in speech and language and has begun to elucidate sensorimotor cortical circuits for speech. This work helps bridge the gap between functional anatomical models of speech and language and models of the functional organization of cortex more generally (Hickok and Poeppel, 2000).

—Gregory Hickok

References


Augmentative and Alternative Communication: General Issues

It is estimated that from 8 to 12 out of every 1000 individuals have a communication disorder severe enough to require the use of augmentative and alternative communication (AAC) intervention (Beukelman and Ansel, 1995). A large percentage of these individuals are children with spoken language disorders and a range of etiologies. Manual signs, communication boards, and computers with voice output have been developed to provide a means by which children with severe spoken language disorders can acquire language and communication skills. AAC is a language intervention approach.

The American Speech-Language-Hearing Association defines AAC as an area of research, clinical, and educational practice that attempts to compensate, either permanently or temporarily, for the impairment and disability patterns of individuals with severe expressive and receptive communication disorders that affect spoken, gestural, and/or written modes of communication. AAC is comprised of a system of four integrated components: symbols, aids, techniques, and strategies. The first component, symbols, are visual, auditory, and/or tactile, used to represent vocabulary and described as either aided or unaided. An aided AAC symbol employs the use of an external medium (e.g., photographs, pictures, line drawings, objects, Braille, or written words), while an unaided AAC symbol utilizes the AAC user’s body (e.g., sign language, eye pointing, vocalizations). Aids are the second component; an aid is an object used to transmit or receive messages and includes, for example, communication boards, speech-generating devices, or computers. A technique, the third component, is the approach or method used for generating or selecting messages as well as the types of displays used to view messages. Messages may be generated or selected via direct selection or scanning. Direct selection permits a child to communicate messages from a large set of options using, for example, manual signing or pointing with a finger or headstick to a symbol. Scanning is used when message choices are presented to the child in a sequence and the child makes his or her selection by linear scanning, row-column scanning, or encoding. Displays may be either fixed (i.e., the symbol remains the same before and after activation) or dynamic (i.e., the symbol visually changes with its selection). Finally, strategies, the fourth component, are the specific intervention approaches in which AAC symbols, aids, and techniques are used to facilitate or develop language and communication skills via AAC (see ASHA, 1991; ASHA, in preparation, for complete definitions).

The role an AAC system plays in a particular child’s life will vary depending on the type and severity of the child’s language disorder. Children who use AAC include those individuals who present with congenital disorders as well as those individuals with an acquired language disorder. Children with congenital language disorders include children with cerebral palsy, dual sensory impairments, developmental apraxia of speech, language learning disabilities, mental retardation, autism, and pervasive developmental disorders. Acquired language disorders may include traumatic brain injury (TBI) and a range of other etiologies (e.g., sickle cell anemia) that affect language skills.

Children with language disorders who can employ AAC systems may range in age from toddlers to adolescents (Romski, Sevcik, and Forrest, 2001). The role AAC plays in language intervention depends on the child’s individual communication needs. It is not restricted to use with children who do not speak at all and may benefit children with limited or unintelligible speech as well as those young children who may be at significant risk for failure to develop spoken communication. There are no exclusionary criteria or prerequisites for learning to use an AAC system (National Joint Committee, in preparation). Every child can communicate! Communication is defined in the broadest sense as “any act by which one person gives to or receives from another person information about that person’s needs, desires, perceptions, knowledge, or affective states” (National Joint Committee, 1992). The modes by which children can communicate range along a representational continuum from symbolic (e.g., spoken words, manual signs, arbitrary visual-graphic symbols, printed words) to iconic (e.g., actual objects, photographs, line drawings, pictographic visual-graphic symbols) to nonsymbolic (e.g., signals such as crying or physical movement) (Sevcik, Romski, and Wilkinson, 1991). AAC interventions incorporate a child’s full communication abilities, including vocalizations, gestures, manual signs, communication boards, and speech-generating devices. Even if a child uses some vocalizations and gestures, AAC systems can augment communication with familiar and unfamiliar partners across multiple environments. Some children with severe spoken language disorders who have no conventional way to communicate may express their communicative wants and needs in socially unacceptable ways, such as through aggressive or destructive, self-stimulatory, and/or perseverative means. AAC systems can replace these unacceptable means with conventional communication (Mirenda, 1997). AAC is truly multimodal, permitting a...
child to use every mode possible to communicate messages and ideas.

While AAC is an intervention approach, a team assessment is needed to describe, within a functional context, the child’s language and communicative strengths and weaknesses and to determine what type of AAC system will permit the child to develop language and communication skills in order to participate in daily activities. AAC assessment is an ongoing process and includes a characterization of the child’s current communication development (i.e., speech comprehension skills, communication modes), an inventory of the child’s environments including partners and opportunities for communication, and a description of the child’s physical abilities to access communication, including vision, hearing, and fine and gross motor skills. Fine and gross motor access includes physical access to an AAC system and in some cases seating and positioning options for optimal communication. A collaborative team approach to AAC service delivery incorporates families and a range of professional disciplines including, though not limited to, speech-language pathologists, general and special educators, and physical and occupational therapists. AAC abilities may change over time, although sometimes very slowly, and thus the AAC system selected for the present may need to be modified as a child grows and develops.

Not surprisingly, standardized psychological and speech and language assessment batteries are often difficult to employ with children with severe spoken language disorders because of the severity of their oral communication impairments. These assessments may not reveal an accurate picture of a child’s abilities since many of these assessments are language-based and may be biased against a child who cannot speak. Often, the children are unable to obtain basal scores on such tests or their scores are so far below those of their chronological age peers that converting a raw score into a standard score is not possible. Systematic behavioral observation within everyday environments and informal measures that inventory and describe communication demands in these settings are employed to measure the communication skills of children who will employ AAC systems rather than standardized tests within isolated settings.

For most children who use AAC systems, language and communication development is the most important goal. Like all language and communication interventions, the long-term goal is to facilitate meaningful communication interactions during daily activities and routines. Goals should not only focus on the technological means of access the child uses, but on the development of language and effective communication skills. Depending on the child’s current language and communication skills, goals may range from developing a basic vocabulary of single symbols or signs to express basic wants and needs to using sentences of symbols and signs to convey complex communicative messages (Reichle, York, and Sigafoss, 1991; Romski and Sevcik, 1996). It is essential that AAC system use take place in inclusive environments. The literature strongly suggests that AAC systems can be embedded effectively within ongoing events of everyday life (Beukelman and Mirenda, 1998).

Using AAC systems in inclusive settings requires that the team work together to ensure that the child has access to his or her AAC device throughout the day and that all adults and children who may interact with the child serve to support the child’s communications as needed.

One frequently asked question is whether the use of AAC systems hinders speech development. Developing natural speech and literacy abilities are extremely important goals of AAC intervention. The empirical evidence suggests that AAC system use may result in increases in vocalizations and in some cases the development of intelligible speech (Beukelman and Mirenda, 1998). There is no evidence to suggest that AAC hinders or halts speech development. The use of AAC systems may also facilitate the development of early literacy skills and later reading.

In summary, for children with severe spoken communication disabilities, the AAC assessment is an ongoing process that includes information about the child’s communication development, the child’s environments, and the child’s physical abilities. Children with severe language disorders who use AAC systems can demonstrate communication achievements far beyond traditional expectations. Recommended assessment and intervention practices are continuing to develop. The use of appropriate AAC systems enables the child to communicate effectively at home, school, play, and work. In addition to the development of communication skills, AAC increases social interactions with family and friends and participation in life activities.

See also AUGMENTATIVE AND ALTERNATIVE COMMUNICATION APPROACHES IN CHILDREN.

—Mary Ann Romski, Rose A. Sevcik, and Melissa Cheslock

References


statement and report on the eligibility of persons with severe disabilities to benefit from communication services and supports.


Further Readings


Bilingualism and Language Impairment

This article is concerned with children who grow up learning two languages simultaneously. These are children who are exposed to two languages on a regular and consistent basis beginning within the first year of birth. (Children who begin learning a second language after 1 year of age are not considered because their pattern of development may be quite different from that of simultaneous bilinguals.) Understanding the nature of and developing appropriate treatment for impairment in simultaneous bilingual acquisition (i.e., bilingual impairment) is of the utmost importance because of the large number of such children worldwide. This article discusses key features of normal bilingual acquisition and factors that can influence bilingual acquisition that do not implicate impairment. It is imperative to understand normal bilingual development if valid diagnosis and treatment of bilingual impairment is to occur. Research on impaired bilingual acquisition and its treatment are then discussed.

What Do We Know About Normal Bilingual Development?

Contrary to the widespread view that simultaneous acquisition of two languages is beyond a child’s normal capacity, research on prenatal and newborn infants indicates that there are no neurocognitive limitations on infants’ innate capacity to acquire two languages simultaneously (Genesee, 2001). Indeed, a growing body of research on children acquiring two languages simultaneously indicates that key milestones in phonological, lexical, syntactic, and pragmatic development occur within the same age range for bilingual children as for monolingual children (Paradis, 2000; Deuchar and Quay, 2000; Comeau and Genesee, 2001; Genesee, 2002a). Of course, there is considerable individual variation in the rate and pattern of normal language development among bilingual children as among monolingual children, and this should be taken into account when identifying possible cases of bilingual impairment. Delay in the emergence of key milestones and variations in pattern of development are not necessarily symptomatic of underlying impairment, although they might warrant careful monitoring.

Phonology. Preverbal bilingual children progress from a stage in which there appears to be no system in either language to distinct phonological patterns in each (Deuchar and Quay, 2000; Paradis, 2000). En route to acquiring the target system, young bilingual children may demonstrate phonological patterns that deviate from those exhibited by monolingual children. The deviations that occur are not necessarily symptomatic of impairment but may simply reflect the bilingual child’s transitional mastery of the complex dual phonological input that the child is exposed to. In the long run, most children exposed to two languages simultaneously and consistently exhibit no phonological difficulties as they mature, as demonstrated by young children’s remarkable ability to acquire native-like accents, in comparison to the notorious phonological disadvantage of older second-language learners.

Vocabulary. Bilingual children generally utter their first words around the same time as monolingual children, and the lexical repertoire of bilingual children is generally of the same magnitude and scope as that of same-age monolingual children when both languages are combined (Genesee and Nicoladis, 1995). When their vocabulary in each language is considered separately, it may be smaller and more restricted in scope than that of same-age monolingual children. Such differences are most likely attributable to the distinct environments in which they acquire each language and usually disappear with age as the child’s experiences in each language expand. It is not uncommon for domain-specific differences in lexical proficiency to persist into adulthood,
however, if the bilingual individual’s contexts for acquiring and using both languages are distinct.

Many bilingual children exhibit “dominance” in one language; this can be reflected in syntactic and pragmatic as well as lexical domains. Dominance can express itself as differential proficiency, preference, or accuracy of use in one language in comparison to the other. In the case of vocabulary, for example, this can result in the child overusing words from the dominant language even in contexts where the nondominant language is appropriate. Dominance is a normal feature of bilingual acquisition and can continue into adulthood, often as a result of greater exposure to one language. Such imbalances do not usually imply impairment. Dominance should be considered carefully when understanding bilingual children’s language development since it can explain their reliance on or more advanced proficiency in one language in comparison to the other.

Syntax. Contrary to earlier views (Volterra and Taeschner, 1978, for example), it is now clear that bilingual children develop separate grammatical systems for each language (Genesee, 2000; Meisel, 2001). This is evident as soon as they begin producing language that is clearly organized according to grammatical principles (in English, from the two-word stage onward). For the most part, bilingual children demonstrate the same stages and patterns of syntactic development in each language as children who acquire the same languages monolingually (e.g., Deuchar and Quay, 2000; Juan-Garau and Pérez-Vidal, 2000). Some bilingual children may show transfer (or so-called interference) effects such that a grammatical pattern (rule) from one language appears inappropriately when the child uses the other language (Döpke, 2000; Yip and Matthews, 2000). Such transfer effects are usually limited in scope and generally reflect grammatical overlap in the two languages. When transfer occurs, it is often, although not always, associated with much greater proficiency in or exposure to one of the languages. Transfer effects are usually short-term, provided the child continues to receive consistent and rich exposure to both languages. Some bilingual children also exhibit more advanced levels of syntactic development in one language than the other (Paradis and Genesee, 1996). This can be due to greater exposure to that language, inherent differences in the acquisition of specific syntactic patterns in the two languages, or simply a preference for one language. These patterns are normal and are not due to impairment.

Pragmatics. There is an extensive body of research on the development of pragmatic (or conversational) skills in bilingual children. Bilingual children are able to use their two languages differentially and appropriately with others, even strangers with whom they have had no or limited contact; this is evident even in children in the one and early two-word stage (Genesee, Nicoladis, and Paradis, 1995; Lanza, 1997; Comeau and Genesee, 2001; Genesee, 2002b). Bilingual children’s pragmatic abilities can be limited by their proficiency. In particular, they may not stick to the language of their conversational partner if their vocabulary, syntactic, or pragmatic skills in that language are not well-developed. In such situations, the child may call up the resources of the other language and code mix. Code mixing is the use of sounds, words, syntax, or pragmatic patterns from both languages in the same utterance or stretch of conversation. Some bilingual children may even prefer to code mix because they are accustomed to using and hearing others use two languages in the same conversation. Indeed, some bilingual children may never have encountered a monolingual person and thus are not used to communicating with adults whose skills are limited to one language. Indeed, code mixing is a normal part of interpersonal communication among fully proficient bilingual adults, and thus young bilingual people are often exposed to proficient adult bilinguals who mix.

Contrary to earlier views, it is now clear that bilingual code mixing is not a sign of linguistic confusion or incompetence (Genesee, Nicoladis, and Paradis, 1995; Meisel, 1989, 2001). To the contrary, child bilingual code mixing, like adult code mixing, is not random but is constrained according to the grammars of the two participating languages (Allen et al., 1999; Paradis, Nicoladis, and Genesee, 2000). In other words, children do not usually violate the grammatical rules of either language when they code mix. Bilingual code mixing in children is also situationally constrained, and bilingual children can adjust their rates of code mixing according to the rates of mixing of their interlocutors (Comeau, Genesee, and Lapagarette, in press). In short, code mixing is a communicative resource that bilingual children use to extend their communicative competence.

Bilingual children’s language usage, including their code mixing, is shaped by the sociocultural context in which they acquire their languages, leading in some cases to patterns that could be misinterpreted. For example, a bilingual child may speak a language variety with phonological, lexical, or grammatical features that would be considered deviant from the point of view of the standard language but are normal in the child’s variety (Crago, 1992). They may exhibit conversational patterns, such as silence, that could be interpreted as lack of pragmatic competence or even language disability from the perspective of mainstream norms but are normal and appropriate in the child’s cultural community (see Crago, 1992, for an example among the Inuit). It is important to consider sociocultural factors as possible explanations of patterns of bilingual usage that might otherwise be attributed to impairment.

What Do We Know About Impairment in Bilingual Acquisition?

It is often thought that children exposed to two languages early in development will experience a higher incidence of impairment than monolingual children and that their impairment is likely to be unique and more severe than that of monolingual children. These expectations are based on the assumption, noted earlier, that
acquisition of two languages simultaneously (or consecutively) during the preschool years exceeds the child’s innate endowment to learn language. In effect, it is thought that bilingualism causes the impairment. However, such an assumption is misguided. While we currently lack adequate normative studies of the incidence of impairment in bilingual children, the extant evidence provides no reason to believe that exposure to two languages causes more delayed or impaired development than one would find in a monolingual population. Moreover, given the overwhelming evidence that bilingual children demonstrate the same milestones and patterns of development as monolingual children, there is no reason to expect unique patterns of impairment among bilingual children. Indeed, Paradis et al. (2003) found that the pattern and severity of impairment in a group of French-English children with a clinical diagnosis of impairment did not differ from that of age- and language-matched monolingual and bilingual controls also with impairment. In contrast, Crutchley and her colleagues found that bilingual children referred to special language units for children with language impairment in Britain exhibited more severe and unique patterns of difficulty on a variety of standardized measures in comparison to their monolingual counterparts (Crutchley, Conti-Ramsden, and Botting, 1997). However, these findings must be interpreted with caution, since, as noted by the authors, there may have been a bias toward inclusion of more severely impaired bilingual children in this sample. As well, the use of norm-referenced tests standardized on monolingual children may bias interpretation toward impairment, since normal bilingual-specific patterns of development were not taken into account. In support of the acquisition results, treatment studies with impaired bilingual children (both simultaneous and consecutive) have demonstrated that outcomes following bilingual treatment are just as positive or even more positive than those following monolingual treatment (Gutierrez-Clellen, 1999; Perozzi and Sanchez, 1992; Thordardottir, Weismer, and Smith, 1997). In sum, and contrary to the bilingualism-as-risk notion, there is no evidence that bilingual impairment is more severe than or different in kind from monolingual impairment, and there is no evidence to support monolingual treatment over bilingual treatment.

On the basis of this evidence, it is recommended that impairment in children acquiring two languages simultaneously be assessed in the same manner as in monolingual children, taking into account what we know about normal bilingual acquisition and the factors that can influence it: exposure, dominance, and sociocultural context. In addition to current best practices in assessment of monolingual children, the following principles should be observed when assessing bilingual children in order to ensure a valid diagnosis of impairment. (1) Evidence for impairment should be attested in both languages. (2) The pattern of impairment in each language should resemble that of monolingual children with impairment acquiring the same languages. (3) Standardized language tests that are normed on monolingual children should not be used normatively, nor should they be the sole basis for the diagnosis of impairment in bilingual children, since the latter may demonstrate normal patterns of performance that could be construed as impaired when compared with monolingual norms (Dodd, So, and Wei, 1996). Standardized tests that are normed exclusively on monolingual children are not likely to make allowances for the sociocultural and exposure differences that bilingual children experience learning two languages, and as a result, they are likely to misrepresent performance differences that bilingual children present as underlying impairment. Decision criteria that recognize alternative paths to normal language development must be used in the diagnosis of impairment in children acquiring two (or more) languages simultaneously.

There is no evidence that language impairment is due to or exacerbated by the simultaneous acquisition of two languages. It is likely that impairment is due to a fundamental problem in the child’s innate capacity to acquire language that manifests itself in whatever language or languages the child is learning. Thus, children in the process of acquiring two languages who are suspected of impairment should not be limited to one language on the assumption that this will benefit their language development. Nor should treatment be restricted to one language only. To the contrary, children who have been exposed to two languages from birth may experience significant personal trauma and sociocultural disadvantages if they are deprived of the benefits of knowing two languages. Moreover, restricting children suspected of impairment to one language may entail significant long-term economic, professional, personal, and social disadvantages in the case of individuals living in bilingual communities. While we still have much to learn about typical and impaired bilingual acquisition, there is no evidence at present that would recommend or justify a decision that bilingual children with impairment learn only one language.

See also BILINGUALISM, SPEECH ISSUES IN.

Acknowledgments

I would like to thank Martha Crago and Elin Thordardottir for helpful suggestions on earlier versions of this chapter. I would also like to thank the Social Sciences and Humanities Research Council, Ottawa, Canada, for support of my research cited in this chapter.

—Fred Genesee

References


Further Readings


Communication Disorders in Adults: Functional Approaches to Aphasia

Functional communication has been a clinical theme since Martha Taylor Sarno first used the term as a label for her Functional Communication Profile (1968). Since then, the concept of functional communication has broadened in scope, with the result that there are now within this field two pertinent connotations for the word functional. Both are applicable to functional approaches to assessment and treatment of communication disorders in adults. Elman and Bernstein-Ellis (1995) suggest that the first connotation invokes a sense of the basics: for example, having the language necessary for signaling survival needs or rudimentary wants, for getting help, or for using “yes” and “no” reliably and accurately. Functional in the second sense connotes smooth running, getting through the worst of one’s communication problem, or learning satisfactory compensatory skills that permit individuals only occasionally to have to remind themselves that they still have residual problems in communicating.

The term communication as used in the assessment and treatment of adult language disorders also has two connotations. For some clinicians, communication is almost synonymous with language, and their work emphasizes recovery or restitution of language skills. But clinicians whose interest is on functional communication utilize a more comprehensive definition. For them, communication typically encompasses not only language, but also other behaviors that permit individuals to exchange information and socialize even when they speak different languages. Most pertinent to adult language disorders are gesturing, drawing, and other ways of getting messages across, or learning how to guide others to provide the support and scaffolding that facilitates interpersonal interchange.

These expanded definitions are crucial to understanding the differences between functional and more traditional approaches to assessment and treatment of the language disorders that are acquired in adulthood, typically the result of insults to the brain and occurring to individuals who previously had normal language and communication. For such individuals, understanding the way that language functions in communication remains relatively spared, in contrast to their deficits of impaired lexicon, grammar, and phonology. As a result, functional approaches tend to stress communication strengths rather than linguistic deficits.

Because they emphasize everyday language and communication use, functional approaches rely heavily on the context in which such activities occur. They focus on authentic interpersonal exchange and interaction across a variety of settings, as well as communicative activities that occur in everyday life. Functional approaches also include usual conversational partners and emphasize their new role in facilitating as normal communication as is possible. With this background in mind, the following summarizes functional approaches to assessment and treatment of aphasia. Although functional approaches can be applied to disorders such as traumatic brain injury and dementia, the bulk of the literature concerns aphasia, and it will be featured here.

Functional Assessment

A recent report from the National Committee on Vital and Health Statistics (2001) foreshadows the emerging emphasis on the need to lay “the groundwork for greater use of functional status information in and beyond clinical care.” This report also supports the World Health Organization’s revised International Classification of Functioning, Disability and Health (ICF, 1999). The ICF makes it clear that in addition to measuring impairment such as aphasia, assessment must also consider how that impairment limits an individual’s ability to go about the activities of daily living. The ICF takes one more step: It also requires the assessment of the effects of activity limitations on the ability to resume one’s previous level of participation in society. Measuring activities and limitations brought about by aphasia, as well as how one’s participation is affected, is precisely the domain of functional assessment. Such assessment does not substitute for tests that inventory the nature and extent of aphasic impairment. However, it dictates that such procedures must be supplemented by other measures.

Functional communication assessment measures are far-ranging. They include observing aphasic persons’ communicative interactions, interviewing aphasic individuals and their families about communication needs, and analyzing their discourse and conversation. A few formal tests, such as Communicative Activities of Daily Living (CADL-2; Holland, Frattali, and Fromm, 1999) and rating scales such as the ASHA FACS (Frattali et al., 1995), the Functional Communication Profile (FCP; Sarno, 1969), and the Communication Effectiveness Inventory (CETI; Lomas et al., 1989) are used to measure activities and activity limitations in ICF terms.

The natural alignment of functional and pragmatic approaches also extends to ICF’s next level, addressing restrictions in societal participation. Being able to resume activities and to participate in society clearly relates to quality of life. There are many quality-of-life measures available, but few at present focus specifically on the effects of communication problems. Nonetheless, effective functional assessment at the level of participation should be considered through interview and observation until more formal measures are available. This is because the ultimate goal of good therapy (functionally oriented or not) is to improve the quality of an aphasic person’s life in this broadest sense.

We now turn our attention to functional approaches to the treatment of adult communication disorders. Specific clinical techniques for functional treatment are methodologically similar to those that are used in more traditional, impairment-focused treatments. That is, principles of learning and counseling are also used in functional treatment approaches. However, there is a great difference in the tasks that comprise traditional
and functional treatments. In the latter, both stimuli and treatment tasks are geared to everyday events and interactions, or to communication strategies that can be used when language skills break down.

An individual’s own pattern, style, and opportunities for communication are emphasized. As a result, the process is less clinician-driven and more the result of collaboration between aphasic individuals, their families, and their clinicians than are most other approaches. Functional approaches encourage the participation of aphasic individuals as well as their families in choosing treatment goals that are almost always cast in everyday terms. The clinician’s role in the goal-setting phase treatment is to guide and counsel about how realistic the goals are, or to propose modifications that might be acceptable. For example, instead of the clinician’s unilateral decision to work on general impairments in word retrieval (which might target words chosen on the basis of their frequency of occurrence or imageability), a collaborative approach might result in work that features retrieving the names of family, friends, and pets. Thus, even if traditional wisdom suggests that treatment should begin with easy targets and proceed from them hierarchically, the functionally motivated clinician might conclude that personal relevance is more important. The Functional Communication Planner (Worrall, 1999) provides a systematic format for determining the focus of such treatment.

There are many variations to the functional treatment theme when it is directed specifically toward the aphasic person in individual treatment. Perhaps the best known functional approaches involve training aphasic individuals to use strategies that facilitate their speech or aid in improving their comprehension. Some examples include supplementing or even substituting speech attempts with drawing or writing, or, for auditory comprehension, teaching aphasic persons to request repetition of a message that he or she does not understand. Often these alternative strategies are practiced in activities that promote the exchange of unknown information in a manner that approximates the normal interchange of everyday communication (PACE treatment; Davis and Wilcox, 1985).

Other approaches involve developing and practicing scenarios and scripts that can be used to recount important aspects of the person’s life, such as how an aphasic man met his wife. A related approach might be to work on specific situations of personal relevance. These situation-specific scripts can be as diverse as a script that aids an individual in getting help in an emergency to plan a trip or even telling a few jokes. Finally, group treatment focused on conversational skills is becoming increasingly more frequent, as illustrated by the rich examples provided in Elman’s book (1999).

There are also a number of ways to train others in the aphasic person’s environment to take a disproportionate share of the burden of communication. Supported conversation (Kagan, 1998) is one such approach. Others include conversational coaching (Holland, 1991), and less formal training, accomplished both through didactic means and by counseling. Partner-centered approaches have been successfully used with family members, volunteers and even clinicians. The approaches share the rationale that communication can be improved not only by improving communication skills in aphasic individuals, but also when others in the communicative environment learn how to listen more effectively, how to encourage multimodal communicative attempts, and how to ask questions of aphasic people more appropriately.

In summary, functional approaches rely on a real-world perspective. Clinicians typically ask, How will this affect this person’s daily life, and that of his or her family? Finally, functions of language, such as being able to invite, deny, or request, are essential to daily social interactions, and therefore to functional approaches.

In the last 35 years, researchers and clinicians have refined their abilities to assess everyday communication and to measure the functional outcomes of treatment efforts. Treatment methods that focus on the functional have also been expanded, and data now support the observation that functional changes can be achieved as a result. Qualitative as well as quantitative research on functional approaches can be expected to expand our knowledge over the next corresponding time period.

—Audrey L. Holland and Jacqueline J. Hinckley

References


Further Readings


Communication Disorders in Infants and Toddlers

Prior to 1986, many speech-language pathologists serving infants and toddlers with communication disorders used an expert service delivery model. In the expert model, the speech-language pathologist is viewed as the professional who provides solutions by way of direct intervention to a child. Service delivery is often direct, and families have little control over the focus of and method of intervention. However, with passage of Public Law 99-457, a shift in service delivery philosophies occurred, with a new emphasis on a family-centered model (Donahue-Kilburg, 1992).

PL 99-457, Part H, mandated comprehensive, coordinated, community-based and family-centered services for infants and toddlers exhibiting disabilities in physical, cognitive, communication, social or emotional, and/or adaptive development. A range of services, including speech-language services and audiology, are available at no cost to the parents except where federal or state law provides for a system of payment by families. Reauthorization of IDEA in 1997 led to a change in name (now referred to as Part C). The reauthorization emphasizes services to children in natural environments (i.e., locations where the child would be served if he or she did not have a disability), using family-directed service delivery to identify family needs, and affirms families as members of the evaluation team.

Interagency coordinating councils (ICCs) exist at the federal, state, and local levels. Eligibility criteria for
early intervention services under Part C is decided by each state's lead ICC, and as a result, varies from state to state. Eligibility is often determined by the presence of a developmental delay in physical, cognitive, speech and language, social or emotional, and adaptive (i.e., self-help) skills; or eligibility may be based on the degree of risk that the child has for developing a delay. There are three types of risks: established risk, biological risk, and environmental risk. In the case of established risk, a child displays a diagnosed medical condition, such as Down syndrome, fragile X syndrome, or Turner's syndrome, that is known to influence development negatively. Children with an established risk qualify for early intervention services. In contrast, a child who is biologically at risk exhibits characteristics (e.g., very low birth weight, otitis media, prematurity) that may result in developmental difficulties. A child with an environmental risk is exposed to conditions that may interfere with normal development, such as poor nutrition, poor environmental stimulation, or caregivers with substance abuse problems. Children with biological or environmental risks are considered to be at risk rather than to have an established risk. Some states include children who are at risk in their eligibility criteria for services; others do not.

One group of toddlers seen by speech-language pathologists who have been studied extensively display slow expressive language development (SELD). SELD is characterized by an expressive vocabulary of less than 50 words at 24 months of age and no word combinations, with no known hearing, cognitive, emotional/behavioral, gross neurological, oral-motor, or environmental deprivation (Rescorla, 1989; Paul, 1993). Paul (1996) concluded that children with SELD should not be regarded as having a disorder; rather, they should be considered at risk for further language impairment. Her longitudinal study of children identified as having SELD indicated that approximately 74% of the children identified with SELD as toddlers were no longer classified as having an expressive language delay by kindergarten age. Paul recommended a “watch-and-see” policy for children with SELD who do not display other risk factors. The watch-and-see policy would monitor children on a regular basis for their linguistic progress. However, other researchers have opposed such a policy, for a variety of reasons (van Kleeck, Gillam, and Davis, 1997). Children with specific language impairment (SLI), which can include receptive and expressive language skills, may also be seen for early intervention services. Risk factors for SLI include heredity, long periods of untreated otitis media, and parental characteristics such as low socioeconomic status, directive interaction style, and extreme parental concern (Olswang, Rodriguez, and Timler, 1998).

The family plays a vital role in providing services to infants and toddlers with communication disorders. They are part of the team from the referral stage through assessment, intervention, and dismissal. It is critical to consider the cultural variables that contribute to each family's unique system of functioning when working with families of infants and toddlers with communication disorders. The speech-language pathologist should understand each family's cultural belief system as it applies to views about disabilities, communication behaviors, and childrearing. Adopting a family-guided approach to early intervention requires a nonjudgmental respect for the family's views. The first step to understanding and respecting family and cultural views is to understand one's own family and culture. This places the speech-language pathologist in a position to understand differences of opinion and reduce potential misunderstandings. The next step is to learn about the family's belief system by observing, listening, and sharing information. Skills in interviewing and counseling are crucial for obtaining and clarifying information in a non-threatening and respectful manner.

Unlike the expert model, which focuses only on the client's needs, the family-guided model focuses on both the family and the child. Child assessment and family assessment may both take place. A family assessment is voluntary and often conducted through interviews and surveys designed to collect information about the family's resources, priorities, and concerns. Child assessments must be multidisciplinary and comprehensive, conducted by trained personnel, and include both strengths and weaknesses of the child. The assessment should also be nondiscriminatory and confidential. Informed consent must be obtained. Professionals have 45 days from screening to complete the evaluation process that may determine the presence of a disorder and determine eligibility.

Assessment of infants and toddlers includes a variety of tasks and a range of informational questions asked of family members. Areas assessed include infant state behaviors, respiration, oral-motor skills (including sucking and swallowing evaluations), nonverbal communication behaviors, play behaviors, caregiver-child interaction, receptive and expressive vocabulary skills, phonological skills, word combinations, syntactic development, functions for language use, and cognitive skills (Donahue-Kilburg, 1992; Dickson, Linder, and Hudson, 1993; Paul, 2001). Assessment in the neonatal intensive care unit may involve working with nurses and parents and revolves around making the environment as positive as possible for the infant (Ziev, 1999). Assessment includes formal, standardized tests, informal methods such as play-based assessment and temptation tasks, and parental report.

Often assessment is conducted in an arena format, where one member of the team serves as the primary facilitator. Other team members observe and make notes about the child's behavior in their specialty area. In a transdisciplinary arena assessment, the primary facilitator will adopt the roles of the other professionals. Families can be involved in assessment by providing information, observing and collecting information, and interacting with their child. The level of involvement is the family's choice. It is important to remember that families are dynamic systems and that roles and functions of family members can change. A family initially
reluctant to participate may want to participate in the next session. Frequent communication is necessary to ensure that services remain family-guided and family-friendly.

Families also assist the speech-language pathologist in determining outcomes rather than goals. These outcomes can be child-oriented, family-oriented, or both. The outcomes are written in family-friendly language, often in the family’s own words. These outcomes are written in the form of an Individual Family Service Plan (IFSP). Importantly, the IFSP is a process as well as a document. The process begins at referral and involves getting to know the family and child. It is an informal transferal of information that builds the foundation of trust and respect underlying future interactions. The written IFSP will contain information about the person who serves as the family service coordinator, the child’s status in the areas of physical, cognitive, communicating, social-emotional, and adaptive development, a plan that indicates the frequency and length of services, who will pay for services, the method of providing services, and how service will be provided in natural environments. The family service provider is selected by the family and is the main coordinator of services. The IFSP also includes a transition plan for children leaving Part C services and entering preschool (or Part B) services. It is optional on the part of the family whether or not they want the IFSP to contain information on the family’s resources, concerns, and priorities. IFSPs can be used for children ages 0–6 years. Thus, on transitioning to Part B services, families may choose an IFSP instead of an Individual Education Plan.

Once outcomes have been selected, the type of intervention approach is determined. It is important to understand the family’s view of speech and language development and the role they believe they play in their child’s development of speech and language. Family members may choose to play an integral role in the intervention; others may prefer to have the intervention conducted by the speech-language pathologist. Intervention studies thus far have provided evidence that early intervention with infants and toddlers can facilitate prelinguistic communication, expressive vocabulary, phonology, social skills, and early word combinations (Girolametto et al., 1996, 1997, 1999; Robertson and Ellis Weismer, 1999; Yoder and Warren, 1998; Loeb and Armstrong, 2001). Some child-oriented techniques include parallel talk, recasting, and expansions. Some interventions focus modeling language effectively in everyday routines and scripts. The ultimate goal of this early intervention for children with established risk and those who are at risk is to reduce the likelihood that they will require future intervention and special education services and will gain the skills needed to participate socially and academically as they grow.

See also SPEECH DISORDERS IN CHILDREN: BIRTH-RELATED RISK FACTORS.

—Diane Frome Loeb

References


Further Readings

Communication Skills of People with Down Syndrome

Down syndrome is the most common genetic disorder in children. The genotype involves an extra copy of the short arm of chromosome 21, either as trisomy (95% of cases), a translocation, or expressed mosaicly. This condition is not inherited and occurs on average in about 1 in 800 live births in the United States. Incidence increases as maternal and paternal age increase. Down syndrome affects almost every system in the body. For example, brain size is smaller in adults though the same size at birth, 50% of these children have significant heart defects requiring surgery, neuronal density in the brain is significantly reduced, middle ear infection persists into adulthood, hypotonia ranges from mild to severe, and cognitive performance ranges from normal performance to severe mental retardation. The remainder of this article will summarize the specific speech, language and communication features associated with this syndrome.

The unique features of speech, language, and hearing ability of children with Down syndrome have been detailed by Miller, Leddy, and Leavitt (1999). First is the frequent hearing loss in infants and children, with more than 75% of young children found to have at least a mild hearing problem at sometime in childhood. These hearing problems can fluctuate, but about one-third of children have recurring problems throughout early childhood that can lead to greater language and speech delay. These results suggest particular attention be directed to monitoring responsiveness to everyday speech and frequent hearing testing through childhood. Second, there are unique verbal language characteristics of persons with Down syndrome. Children experience slower development of language relative to other cognitive skills. Communication performance is characterized by better language comprehension than production. Vocabulary use is better than the mastery of the grammar of the language. Progress in speech and language performance is linked to several related factors, including hearing status, speech-motor function status, and advancing cognitive skills associated with a stimulating verbal and nonverbal environment. Progress in speech, language, and communication should be expected beyond early childhood through adolescents (Chapman, Hesketh, and Kistler, 2002). A third unique feature is a protracted period of unintelligible speech. Speech intelligibility is a persistent problem of persons with Down syndrome through late childhood. Most family members have some difficulty understanding the speech of their children in everyday communication. Treatment protocols can improve speech intelligibility, leading to improved communication of children and adults. Finally, the development of writing and literacy skills in persons with Down syndrome should be expected. Children participating in early reading and literacy experiences experience better communication and academic skills than their peers with less literacy experience.
Early reading programs have been successful at teaching sight word vocabulary to children as young as 3 years of age.

**Assessment Principles**

The following principles have evolved from our research over the past 10 years. Children with Down syndrome are very challenging to evaluate. The first and perhaps the only principle is to not make any assumptions about perceptual, motor, and cognitive skills or the child’s experience with oral and written language. We suggest the following as guidelines for developing an assessment protocol. Access all information sources about current communication abilities across contexts—school, home, day care, and community. Review all data available on motor and cognitive development as well as perceptual (hearing and vision) status to direct the assessment decision-making process. Use flexible communication assessment protocols that can meet the specific attention shifts and motivational challenges of people with Down syndrome. Make sure the context of the assessment matches the child’s performance level (i.e., play-based, observation, and standardized measures). Contrast measures conducted within familiar contexts, child- or family-centered approaches to assessment, with those taken in the absence of relevant context, i.e., standardized tests. Assess the child’s communication environments as well as the child’s independence in activities of daily living. Keep in mind that many persons with Down syndrome have a history of failing tests, “escaping” boring formal assessment procedures. Cover all bases evaluating cognition performance, hearing, verbal and printed language comprehension and production, oral nonspeech function, and speech behaviors. Remember that a child’s performance in the office or clinic may represent a very thin slice of the full range of his or her capability.

**Limitations of Standardized Tests Relevant for Children with Down Syndrome**

Coggins (1998) identifies a number of limitations of standardized measures when testing children with developmental disabilities. Most tests are examiner directed, limiting the child’s initiations and spontaneous language to single words or phrases. It is often difficult to translate test results into clinically useful outcomes. Rigid test administration guidelines make it difficult to use most measures with atypical children. Parents are typically excluded from active participation in the assessment process.

**Challenges for Accurate Assessment**

**Consistency of Responding.** One of the most frustrating characteristics of these children is the lack of consistency of responding in assessment tasks. This variability is associated with two things in our experience. The first is rapid shifts in attention and the second is motivation. Clearly, these constructs are interrelated and it would be impossible to determine which is causal for any specific behavior. We have found that motivation is central to maintaining consistent response patterns. If we can provide a task that is sufficiently motivating, attention is maintained and responses are more consistent. Our work suggests that successful assessments can be conducted by careful preparation and understanding the interests of each child, what activities they like, what holds their attention at home and school, and then selecting assessment materials that can be imbedded into these activities.

**Memory.** The work of Michael Marcell (Marcell and Weeks, 1988) documents verbal short-term memory deficits. This has significant implications for assessment of language comprehension and production, particularly when using standardized procedures that require processing specific stimuli and remembering it long enough to provide to appropriate response. Clearly, memory deficits may also be contributing to behaviors that may be labeled as inattention or that result in inconsistent response patterns. In our experience, providing visual support enhances performance when verbal abilities are tested. This may involve pictures, graphic material, or printed words.

**Motor Limitations.** It has been widely reported that children with Down syndrome have motor deficits. Hypotonia is frequently cited as a cause, but there are little data to support this claim. Motor deficits are quite variable, with some children performing at age level and others show significant motor limitations delaying the onset of ambulation and other motor milestones. Testing protocols must take into consideration the motor demands on the child relative to the child’s motor abilities. Make sure that the assessment tasks require motor responses within the child’s capabilities.

**Vision.** France (1992) provides a detailed account of the visual deficits of children with Down syndrome. He followed a group of 90 children and reported that 49% had visual acuity deficits, with myopia being the most common. He also documented oculomotor imbalance in over 40% of the children, convergent strabismus accounting for the majority of these cases. In the majority of these cases only glasses were required to achieve normal vision. The message here is to make sure the children can see the stimuli during testing.

**Hearing.** Hearing remains an issue for children with Down syndrome because of frequent episodes of otitis media. Monitoring hearing should be done every 6 months for the first 10 years of life. In our work we have found that 33% of our children always had a hearing loss, 33% had a loss sometimes, and only 33% never had a loss. This was after screening out all of those children with significant hearing loss due to other causes. When oral language is tested, it is important to know the child’s hearing status on the day of testing.
Summary. Designing a testing protocol requires attention to the skills and abilities the child is expected to bring to the task. These include attention and motivation differences, memory deficits, hearing and visual deficits, and motor limitations. Each of these can compromise the outcome of the assessment if accommodations are not made. It is also clear that in order to optimize the consistency of responding, alternative testing formats will have to be implemented. These testing formats will need to be less rigid, be context-based, and be child-centered rather than examiner-centered. A skilled clinician will have to follow the child’s lead to implement functional, criterion reference, play-based assessments. Observational methods will also provide important information.

Who Is Responsible for the Development of Communication Skills?

The answer is that everyone is responsible—parents, teachers, and speech-language pathologists—but parents have the most pervasive role in the process. A recent book by Betty Hart and Todd Risley documents the contribution parents make to their children’s language development in the first 3 years of life (Hart and Risley, 1995). Their results support what we have known for some time about language development: (1) parents are the first teachers of speech, language, and communication skills; (2) children will follow their parent’s model of communication action and style; and (3) social relations are central for the development of communication skills.

Promoting Family Communication Supporting Language Development

The language and cognitive development of typical children can be improved by simply talking to them, by producing more and longer utterances with more complex vocabulary. This strategy can be used if we keep in mind children’s ability to comprehend the language addressed to them and how parents adjust their own language to optimize the child’s chances of understanding the message. Parents automatically adjust their language to children on almost every linguistic dimension, paying attention to the form of message. They make sure hearing test results are considered and follow the child’s lead. Help maintain the child’s focus (topic and meaning) with your responses, comments, and questions. (4) Value the child. Recognize the child’s partner has something to say worth listening to, so pay attention. (5) Don’t play the fool. A valued conversational partner has something to say worth listening to, so pay attention. (6) Learn to think like a child. Consider that the child’s perspective of the world is different at different levels of cognitive development.

Research on language learning in children with Down syndrome has documented that language learning is occurring and continues through adolescence (Chapman et al., 2002). The recent research on family communication style and frequency of communication may account for why some children with Down syndrome learn language more rapidly than other children. In our experience, families with children making good progress with their language and communication skills share common features. They select language levels relative to their child’s ability to understand the message and not at their ability to produce messages. They have realistic communication goals. These families expect that their child will learn to read. They focus on understanding the content of their child’s message and are not as concerned with the form of message. They make sure hearing testing is scheduled every six months through the developmental period. And they plan frequent outings to provide their children with varied experiences outside the home.

Who Is Running the Show, Parents Versus Professionals?

Diane Crutcher (1993) has powerfully presented the key issues underlying tensions between parents and professionals in speech and language. She articulates three issues parents perceive as limitations of speech-language intervention. The first is the lack of professional time, awareness, or unwillingness to explore intervention techniques specifically for their individual child. Second is the failure to modify intervention techniques into strategies that fit a family’s natural lifestyle. The time constraints that most clinicians work under promote a “one size fits all” mentality. A lack of sensitivity to in-
Dementia

Dementia is a syndrome characterized by deficits in multiple cognitive domains, including short- and long-term memory and at least one of the following: aphasia, apraxia, agnosia, and impaired executive function (American Psychiatric Association, 1994). The deficits must be sufficiently potent to affect social and occupational functioning and apparent in the absence of delirium. Dementia is associated with many disorders, but most commonly with Alzheimer’s disease (AD) and Lewy body disease (LBD). Other common dementia-producing diseases are vascular disease and Parkinson’s disease (PD).

Because communicative functioning is a manifestation of cognition, it is necessarily affected when an individual experiences the multiple cognitive deficits that define the syndrome of dementia. However, differences in the nature and distribution of neuropathology in the common dementia-associated diseases produce unique patterns of communication deficits.

Effect of Alzheimer’s Neuropathology on Communicative Function

The neurochemical alterations, neurofibrillary tangles, and neuritic plaques that characterize AD begin in the entorhinal cortex, perforant pathway, and hippocampal formation. Gradually, cells throughout the neocortex are affected, especially those in temporoparietal cortex. The formation of tangles and plaques eventuates in cell death and interferes with intercellular transmission. Subcortical structures and the motor strip are relatively free of neuropathology throughout much of the disease, which accounts for the fact that the speech of individuals with AD is spared.

Because AD begins in the hippocampal complex, an area important for the formation of recent or episodic memory, the typical initial manifestation is loss of memory for recent events. With disease progression to frontal cortex and temporoparietal association areas, other declarative memory systems are affected, specifically semantic and lexical memory. Because the basal ganglia and motor cortex are spared throughout most of the disease course, procedural memory also is spared.

In the early stages of AD, communicative functions dependent on recent memory, such as holding a conversation, are affected. Affected individuals forget what they have just heard, seen, or said. Many sentences are left unfinished, with forgotten communicative intentions, and repetitiousness is common. Comprehension of written materials, particularly long passages, diminishes because of memory impairment. The mechanics of reading are spared, however, and individuals with mild AD can still write, though they make frequent spelling errors.

The expression of grammar and syntax is remarkably intact, although occasional errors may be made. Individuals with mild AD can usually follow three-stage commands, answer comparative questions, name
familiar items on confrontation, generate exemplars in a category, define familiar words, and describe pictures, although their descriptions do not contain as much factual information as those of age-matched healthy elders (Bayles and Tomoeda, 1993; Hopper, Bayles, and Kim, 2001).

By the middle stages of AD, when affected individuals have become disoriented for time and place and memory problems are more florid, communication is more disrupted (Bayles and Tomoeda, 1995). Meaningful verbal output diminishes because individuals have increasing trouble generating a series of meaningful ideas. Writing words to dictation may remain, but writing letters or pieces of any length is problematic. The ability to read is retained, although affected individuals rapidly forget what they have read. Grammar and syntax continue to resist prominent disease effects.

Persons with mid-stage AD can greet, name, and express many needs. Most can participate in short conversations, especially if those conversations involve only two people; however, they frequently have trouble retrieving desired names. They can answer questions and understand common gestures. Two-stage commands are comprehensible by most persons with mid-stage AD, and some can follow three-stage commands. Reading comprehension for single words remains good. Most individuals with mid-stage disease can still name on confrontation and produce exemplars in a category, but not as efficiently or accurately as normal elders. Verbal output continues to diminish in terms of meaningfulness, and sentence fragments are more common.

By the end stage of the disease, memory loss is extensive, disorientation may extend to self as well as time and place, and problem-solving skills are minimal. Urinary and then fecal incontinence develop, and ultimately ambulatory ability is severely compromised or lost. Nonetheless, a few linguistic skills are intact (Bayles et al., 2000). Most individuals retain some functional vocabulary, although a small percentage are mute. Much of the language produced by those who still speak is nonsensical. Nonetheless, many patients can follow a one-stage command demonstrating comprehension of language. The majority can read a simple word. Many retain common social phrases such as “I don’t care” and “I don’t know,” and can contribute to a conversation.

Effects of Lewy Body Pathology on Communication Function

Lewy bodies are spherical, intracytoplasmic neuronal inclusions that have a dense hyaline core and a halo of radiating filaments composed of proteins containing ubiquitin and associated enzymes (McKeith and O’Brien, 1999). They were first described in the literature by the German neuropathologist Friederich Lewy in 1912. Lewy bodies are classically associated with Parkinson’s disease, particularly in the basal ganglia, brainstem, and diencephalic nuclei. They may also be widespread in the cerebral cortex. Diffuse distribution of Lewy bodies is associated with dementia, and 10%–15% of cases of dementia (Cummings and Benson, 1992; McKeith et al., 1992) have this cause. Patients with LBD often have concomitant Alzheimer’s pathology, and some have proposed that LBD is a variant of AD. However, there are cases of pure LBD, which argues for the theory that LBD is neuropathologically distinct (see Cercy and Bylsma, 1997, for a review).

Symptoms of LBD include fluctuating cognition in 80%–90% of patients (Byrne et al., 1989; McKeith et al., 1992), visual or auditory hallucinations, paranoid delusions, extrapyramidal features, and confusion (McKeith and O’Brien, 1999). Intellectual deterioration is more rapid than that observed in AD patients and disease duration is shorter.

Like the dementia of AD, LB dementia has an insidious onset and progressive course (Hansen, Salmon, and Galasko, 1990), producing the temporoparietal features of aphasia, apraxia, and agnosia (Byrne et al., 1989). Because the extrapyramidal features may be later occurring, LBD in its early stages may be misdiagnosed as AD. Prominent memory deficits may not be the presenting feature but may appear as the disease progresses.

The literature on the effects of LBD on language and communicative functioning is scant. However, it is reasonable to expect that the associated confusion, memory and attentional deficits will disrupt communication in much the same way that they do in AD. Because fluctuating cognitive status is a prominent feature of LBD, clinicians can expect wide fluctuation in communication skills.

Heyman and colleagues (1999) compared the neuropsychological test performance of individuals with AD with that of individuals with AD plus LBD. Both groups of patients were severely impaired on measures of mental status, verbal fluency, confrontation naming, concentration, visuospatial/constructional abilities, constructional praxis, and word list learning and recognition. The sole significant difference between the two groups appeared on the delayed recall of a word list, with AD-only patients being more impaired.

Vascular Dementia

Dementia can be produced by disease of the brain’s vascular system, and the distribution of disease defines the nature of the neuropsychological deficits. However, the many possible variations in the nature and distribution of vascular disease make it impossible to specify a typical profile of neuropsychological deficits. For example, individuals with an infarct in the territory of the left posterior cerebral artery involving the temporoparietal region might exhibit amnesia, whereas an individual with an infarcts in the left anterior cerebral artery involving the medial frontal region may have prominent executive dysfunction, with both nonetheless meeting the criteria for dementia. Similarly, individuals with cortical infarct differ from those with subcortical infarcts.

Individuals at greatest risk for developing vascular dementia are those who have experienced one or more clinically evident ischemic strokes (Desmond et al.,
1999). In fact, one-fourth to one-third develop dementia within 3 months (Pohjasvaara et al., 1998; Tattemichi et al., 1992). Examination of neuropsychological abilities has revealed inconsistent patterns of strengths and weaknesses (Reichman et al., 1991), but executive functions often are disproportionately impaired, as are motor aspects of language production (Powell et al., 1988).

**Parkinson’s Dementia**

Parkinson’s disease is associated with a loss of striatal dopaminergic neurons, particularly in the pars compacta region of the substantia nigra. Tremor is the best recognized symptom and is present in approximately half of individuals with PD (Martin et al., 1983). Often tremor begins unilaterally, increasing with stress and disappearing in sleep. Other early symptoms include aching, paresthesias, and numbness and tingling on one side of the body that ultimately spread to the other side. Other classic motor symptoms are rigidity, slowness of movement, and alterations in posture.

Not all individuals with PD develop dementia, and prevalence estimates vary. Marttila and Rinne (1976), in one of the most comprehensive studies of prevalence, reported it to be 29%. Other investigators have reported similar estimates of dementia prevalence (Rajput and Rozdilsky, 1975; Mindham, Ahmed, and Clough, 1982; Huber, Shuttleworth, and Christy, 1989). Widely debated is the cause of the dementia, with some attributing it to cortical degeneration and others to subcortical damage that impairs neurological control of attention (Brown and Marsden, 1988). Rinne and colleagues (2000) argue that reduced fluorodopa uptake in the caudate nucleus and frontal cortex produces impaired performance on neuropsychological tests that require executive function.

Individuals with PD, regardless of whether they develop dementia, have speech motor deficits because the disease damages the basal ganglia and striatal-cortical circuitry, which are involved in motor function. Those who develop dementia have problems communicating for other reasons, namely deficits in memory, attention, and executive functions. However, considerable evidence exists that language knowledge generally is preserved (Pirozzolo et al., 1982; Bayles and Tomoeda, 1983; Huber et al., 1986). Bayles (1997) argued that impaired performance on tests that manipulate language, such as confrontation naming and sentence comprehension, result more from nonlinguistic cognitive deficits than a loss of linguistic knowledge.

See also Alzheimer’s disease.

—Kathryn A. Bayles

**References**


**Further Readings**


**Dialect Speakers**

A dialect refers to any variety of language that is shared by a group of speakers. It is not possible to speak a language without also speaking a dialect (Wolfram and Schilling-Estes, 1998). Although all dialects of a language are equally systematic and complex, on a social level, dialects are often described as falling on a continuum of standardness. The most standard dialect of a language generally reflects an idealized prestige form that is rarely spoken by anyone in practice. Rules for producing this standard, however, can be found in formal grammar guides and dictionaries. Versions of the standard can also be found in formal texts that have been written by established writers. Next in standardness are a number of formal and informal oral dialects. These dialects reflect the language patterns of actual speakers. Norms of acceptability for these dialects vary as a function of the regional and social characteristics of different communities and of different speakers within these communities. Nonstandard dialects represent the other end of the continuum. These dialects also reflect spoken language, but they include socially stigmatized linguistic structures. Other terms used to describe nonstandard dialects are nonmainstream and vernacular.

At the linguistic level, scholars repeatedly highlight the arbitrary nature of a dialect’s social acceptability
(i.e., standardness). In fact, Milroy and Milroy (2000) argue that contradictory and changing attitudes to the same linguistic phenomenon can emerge at different times in the history of a language. For example, as these authors note, before World War II, absence of postvocalic /r/ in words such as car and park was not stigmatized in New York City. By 1966, however, r-lessness had become a stigmatized marker of casual style and lower social class. English dialects containing r-lessness continue to be stigmatized in the United States, but in England, English dialects with this same linguistic pattern have high status.

Most linguistic patterns that occur in the standard dialects of a language also occur in those that are nonstandard. Seymour, Bland-Stewart, and Green (1998) refer to these language patterns as noncontrastive, because all dialects of a language are thought to share these forms. Despite the similarity that exists among dialects, nonstandard versions are typically described by listing only those language patterns that do not appear in the standard varieties. Seymour, Bland-Stewart, and Green refer to these patterns as the contrastive features. Descriptions of nonstandard dialects become even narrower when they are generated by the media and general public, because these descriptions tend to highlight only the contrastive patterns that are highly stigmatized (Rickford and Rickford, 2000). Zero marking of the copula be (i.e., he walking) is an example of a pattern that is frequently showcased for African-American English (AAE) (Rickford, 1999). Other stereotypic patterns include the use of ya’ll and fixing to to describe versions of Southern White English (SWE), a prefixing (e.g., he was a-walking ...) to characterize Appalachian English, and pronouncing think and that as tink and dat to depict Cajun English.

Although it is relatively easy to identify the language forms that differentiate a nonstandard dialect from one that is viewed as standard, it is much more difficult to identify patterns that distinguish one nonstandard dialect of a language from another. One reason for this is that many nonstandard dialects of a language share the same contrastive patterns. Unique contrastive patterns for different nonstandard dialects are particularly rare when the dialects being compared are produced in the same community and by speakers of the same social class. Oetting and McDonald (2001, 2002) illustrated this finding by comparing the contrastive patterns of two nonstandard dialects spoken in southeastern Louisiana. The data for this comparison were language samples of children who lived in the same rural community and attended the same schools. Forty of the children were African American and spoke a southern rural version of AAE, and 53 were white and spoke a rural version of SWE. The AAE and SWE dialects spoken by the children were deemed distinct through the use of a listener judgment task and a discriminant function analysis of 35 different nonstandard (i.e., contrastive) language patterns found in the transcripts. Nevertheless, of the 35 contrastive patterns examined, 31 of them were found in the conversational speech of both the AAE and SWE child speakers.

Besides contrastive forms, there are three other ways in which dialects differ from one another. One way is in the frequency with which particular language forms are produced. Wolfram’s (1986) analysis of consonant cluster reduction in 11 different English dialects is useful for illustrating this finding. His data showed cluster reduction occurring 3% of the time for standard English and northern white working-class English, 4% for northern African-American working-class English, 10% for southern white working-class English, 36% for southern African-American working-class English, 5% for Appalachian working-class English, 10% for Italian-American working-class English, 22%–23% for Chicano working-class and Puerto Rican working-class (New York City) English, 60% for Vietnamese English, and 81% for Native American Puebloan English. For each dialect listed, the percentage reflects the degree to which consonant clusters were reduced in regular past tense contexts that were followed by a nonconsonant (e.g., ‘Tom live in’). As can be seen, all 11 of the dialects showed cluster reduction, but the frequency with which this pattern occurred in each dialect greatly varied.

A second way in which dialects differ from one another is with the linguistic environments in which particular language forms occur. The effects of linguistic context on language use are often described as linguistic constraints or linguistic conditions (Chambers, 1995). Studies of linguistic constraints typically occur with language forms that show systematic variability in their surface structure. An example of a variable form is one that can be overtly expressed (i.e., present in the surface grammar; he is walking) in some contexts but is zero-marked (i.e., absent from the surface grammar; he walking) in others.

One of the most widely studied variable forms is the copula be, and research on this structure has typically involved nonstandard versions of AAE. At least six different linguistic constraints have been found to influence AAE speakers’ use of the copula be. These constraints include the type of preceding noun phrase (noun phrase versus personal pronoun versus other pronoun), the phonological characteristics of the preceding environment (vowel versus consonant; voiced versus unvoiced consonant), the person, number, and tense of the verb context (first, second, third; present versus past), the grammatical function of the be form (copula versus auxiliary), the nature of the following predicate clause (locative versus adjective versus noun), and the phonological characteristics of the following environment (vowel versus consonant) (Rickford, 1999). The person, number, and tense of the verb context, the grammatical function of the be form, and the nature of the following predicate clause also have been shown to affect copula be marking in various nonstandard SWE dialects (Wolfram, 1974; Wynn, Eyles, and Oetting, 2000). Other morphosyntactic structures of English that have been found to be influenced by linguistic constraints include, but are not limited to, negation, do support, verb agreement, relative pronouns, plural marking, and question inversion (Poplack, 2000).
A third way in which dialects differ from one another is in the semantic meanings or grammatical entailments of some forms (Labov, 1998). These particular cases involve language patterns that occur in most dialects of a language, but their meanings or use in the grammar are unique to a particular dialect. These patterns are often described as camouflaged forms because, on the surface, the contrastive nature of these forms can be difficult to notice (Wolfram and Schilling-Estes, 1998). Use of had + V-ed is an example of a camouflaged pattern. In most dialects of English, this structure carries past perfect meaning (e.g., “I already had eaten the ice cream when she offered the pie”). In some English dialects such as AAE, however, had + V-ed also can express preterite (i.e., simple past) meaning (Rickford and Rafal, 1996).

Ross, Oetting, and Stapleton (in press) provide the following example from a 6-year-old AAE speaker to illustrate the preterite meaning of this form: “Then my mama said, ‘It’s your mama. Let me talk to your daddy.’ Then she had told my daddy to come with us and bring a big rope so they could pull the car home.”

Dialect use is affected by factors that are both internal and external to a speaker (Milroy, 1987; Chambers, 1995; Wolfram and Schilling-Estes, 1998). Internal factors that have been shown to influence the type and density of one’s dialect include age, sex, race, region of the country, socioeconomic status, type of community, and type of social network. Interestingly, regardless of race, region, community, and network, members of lower social classes produce a greater frequency of contrastive dialect forms than members of higher classes. Greater frequencies of contrastive patterns also have been found for younger adults than for older adults, and for males than for females. Exceptions to these generalities do exist, however. For example, Dubois and Horvath (1998) documented a V-shaped age pattern rather than a linear one in their study of Cajun English. They also found that the type and degree of the age pattern (V-shaped versus linear) depended on the speaker’s sex and type of social network. In particular, the V-shaped pattern was more pronounced for men than for women, and only women from closed social networks showed the V-shaped pattern. Women in open networks showed a linear pattern. Interestingly, though, the linear trend reflected higher frequencies of nonstandard dialect use by the older women than younger women. This finding contrasts with what is typically reported in other non-standard dialect work, namely, that older adults present fewer instances of nonstandard forms than younger adults.

Some of the external factors that have been shown to affect dialect use include the type of speaking style (casual versus formal; interview versus conversation), speaking partner (familiar versus unfamiliar; with authority versus without), modality of expression (speaking versus writing versus reading), genre (persuasive versus informative versus imaginative), and type of speech act (comment versus request for information) (for data, see Farr-Whitman, 1981; Labov, 1982; Smitherman, 1992; Lucas and Borders, 1994). Influences of these external factors on dialect use interact in complex ways with those that are internal to a speaker. The dynamic interactions that occur between and among these variables help explain why dialect use is often described as fluid, flexible, and constantly changing.

The challenge for scientists in communication disorders is to learn how a speech or language disorder affects a person’s use of language, regardless of the dialect spoken. Thus far, most descriptions of childhood language impairment have been made within the context of standard dialect varieties only. Extending the study of childhood language impairment to different nonstandard dialects is a topic of current scholarly work and debate (see DIALECT VERSUS DISORDER).

—Janna B. Oetting

References


**Further Readings**


---

**Dialect Versus Disorder**

In 1983, the American Speech-Language-Hearing Association (ASHA) published a position statement on the topic of social dialects. A major point of the publication was to formally recognize the difference between language variation that is caused by normal linguistic processes (i.e., dialects) and variation that is caused by an atypical or disordered language system (i.e., language impairment). Through the position statement, ASHA also formally rebuked the practice of diagnosing and treating any dialect of a language as an impairment. The data used to support this stance were sociolinguistic findings regarding the systematic and complex nature of all dialects, including those that are socially stigmatized (see DIALECT SPEAKERS).

ASHA’s position statement remains relevant today. Research on children’s acquisition and use of different dialects is still a relatively new area of scientific endeavor, but small strides have been made. For example, there now exist numerous articles describing different dialects of English, and current publications about childhood language development, assessment, and treatment now routinely include discussions of dialect diversity. The different types of testing biases that may surface when the language tester and testee come from different linguistic or cultural backgrounds have also been described (Fagundes et al., 1998; Wilson, Wilson, and Coleman, 2000). A few traditional speech and language tests have added alternative scoring procedures for some nonstandard dialects (Garn-Nunn and Perkins, 1999; Ryner, Kelly, and Krueger, 1999). Alternative or new dialect scoring methods also have been created for the analysis of children’s conversational language samples (Nelson, 1991; Stockman, 1996; McGregor et al., 1997; Seymour, Bland-Stewart, and Green, 1998).

Most of the advances listed started with methods and materials that were designed for speakers of standard English. Two different types of changes were then made to these methods. One change involved broadening the range of language forms that are considered normal by including the contrastive patterns of different dialects. The other type of change was to restrict the analysis to only the noncontrastive patterns. Contrastive patterns are those that show variation in surface structure across different dialects of a language (Seymour, Bland-Stewart, and Green, 1998). In English, one contrastive pattern is the copula/auxiliary be. In standard dialects of English, overt marking of be is obligatory in utterances such as “you are walking.” In other dialects, such as African-American English (AAE) and Southern White English (SWE), overt marking of be in this context is optional. As a result, “you walking” and “you are walking” are both felicitous in these dialects. The use of contrast analysis when working with language sample data is an example of an alternative assessment method that treats the contrastive patterns of different dialects as normal (McGregor et al., 1997). With this method, the only language patterns that can be viewed as errors are those that cannot occur in the child’s dialect.

Noncontrastive patterns are those that do not show surface variation across different dialects of a language (Seymour, Bland-Stewart, and Green, 1998). One noncontrastive pattern of English is S-V-O word order (Martin and Wolfram, 1998). This pattern is thought to be noncontrastive because all dialects of English thus far have been shown to present this word order. Other patterns thought to be noncontrastive in English include various forms of complex syntax that make sentential coordination and subordination possible. An example of...
a language assessment method that restricts the analysis to the noncontrastive patterns of dialects is Stockman's (1996) Minimal Competency Core (MCC) analysis. The goal of MCC analysis is to identify and then evaluate a common core of language that can be found in multiple dialects of a language. As a criterion-referenced procedure, MCC specifies a minimum level of competency for each language pattern and each age level examined. One of the language items included in MCC is a child's mean length of utterance (MLU). For English-speaking 3-year-olds, Stockman (1996) sets the minimum MLU at 3.27 morphemes. She also lists 15 consonants in the initial position and a set of semantic expressions and pragmatic functions that all 3-year-olds should be able to demonstrate, regardless of the English dialect they use. Another example of a relatively new assessment tool that targets the noncontrastive features of dialects is that formulated by Craig, Washington, and Thompson-Porter (1998b), which uses Wh-questions and passive probes.

All of these advances treat the contrastive patterns of dialects as problematic for diagnostic purposes. The problem, as articulated by Seymour, Bland-Stewart, and Green (1998), is that some contrastive patterns of some nonstandard English dialects can look very similar to those that are produced by standard English-speaking children who have a language impairment. Some of the surface patterns that are generated by both language learning conditions include zero marking of be (e.g., “you walking”), zero marking of past tense (e.g., “yesterday she fall”), and zero marking of third person (e.g., “today he walk”). Seymour, Bland-Stewart, and Green refer to these patterns and other contrastive forms as presenting a diagnostic conundrum because interpretations of their use as markers of either a normal dialect or a grammatical impairment are difficult. These authors also state that the exclusion of the contrastive patterns within assessment is necessary only until more is known about children’s acquisition and use of these patterns. As more research is completed, new methods that include the contrastive patterns should be made possible.

Recently, research has begun to focus on children’s acquisition and use of the contrastive patterns of dialects. Findings from some of these studies suggest that these particular language forms may not be as problematic as they first seemed. For example, at least four studies have examined the effect of these patterns on standard calculations of children’s average utterance length and utterance complexity. Each of these studies has shown children’s use of the contrastive patterns to play a minimal role within these calculations. For three of the studies, the focus has been on the contrastive patterns of AAE (Craig, Washington, and Thompson-Porter, 1998a; Jackson and Roberts, 2001; Smith, Lee, and McDade, 2001). The participants in these studies have ranged in age from 3 to 9 years. Measures of length have been calculated on utterances, C-units, and T-units, and measures of complexity have involved counts of complex syntax. In every case, children’s use of the contrastive patterns of AAE has been shown to be relatively unrelated to the length and complexity indices. For example, Jackson and Roberts (2001) report correlations between children’s use of contrastive AAE patterns and their utterance length and complexity scores to be at or below -.11.

Oetting, Cantrell, and Horohov (1999) also examined the effect of contrastive dialect forms on standard calculations of utterance length and utterance complexity. The participants in their study were children who spoke a rural Louisiana version of SWE, and they ranged in age from 4 to 6 years. Approximately one-third of the children were classified as specifically language impaired (SLI); the others were classified as normal. Three language indices, MLU, developmental sentence score (DSS), and Index of Productive Syntax (IPSyn), were evaluated. To examine the effect of the contrastive forms, scores for MLU, DSS, and IPSyn were calculated twice for each child, once using samples that contained utterances with contrastive forms and once using the same samples with the contrastive utterances removed. Results indicated that the diagnostic classification of each child as either normal or SLI was the same, regardless of whether utterances with the contrastive forms were included or excluded.

Recent studies also suggest that grammatical weaknesses of children with SLI can be identified within the contrastive forms. In addition to evaluating indices of utterance length and complexity, Oetting, Cantrell, and Horohov (1999) examined the grammatical profile of SLI within the context of nine contrastive forms of SWE. For five of the forms (i.e., third person regular, contractible copula, contractible auxiliary, uncontractible auxiliary, and auxiliary does), the children with SLI were found to present rates of overt marking that were lower than those of their SWE-speaking age-matched and language-matched peers. In a second study, Oetting and McDonald (2001) examined the grammatical weaknesses of SLI in the context of two nonstandard dialects, SWE and a rural Louisiana version of AAE. In this study, 35 different contrastive patterns were coded. Differences between the normal children and those with SLI were identified for 14 of the contrastive patterns. A full model discriminant function that involved counts of all 35 patterns resulted in 90% of the children being correctly classified as either normal or impaired. Stepwise analyses yielded slightly different discriminant functions for identifying children with SLI in the SWE dialect as compared to the AAE dialect, but both models included language forms needed to formulate questions and mark tense. The finding that both dialect groups with SLI were shown to have trouble with these two areas of grammar is consistent with SLI studies that have been completed with standard English speakers, nonstandard English speakers, and speakers of languages other than English (e.g., Rice, Wexler, and Hershberger, 1998; Seymour, Bland-Stewart, and Green, 1998; Craig and Washington, 2000; Paradis and Crago, 2000).

A few recent studies also have examined the developmental trajectories of particular contrastive patterns...
in isolation (Wyatt, 1996; Henry et al., 1997; Jackson, 1998; Burns et al., 1999; Wynn, Eyles, and Oetting, 2000; Ross, Oetting, and Stapleton, in press). The contrastive patterns examined in these studies have included aspectual be and preterite had + Ved in AAE, copula be in AAE and SWE, and negative concord in Bristol English and Belfast English. Each of these studies has shown normally developing children to be remarkably capable of learning the distributional properties of their native dialect. This finding occurs even when children who speak different dialects of a language live in the same community and attend the same schools (Wynn, Eyles, and Oetting, 2000; Ross, Oetting, and Stapleton, in press). Of the studies listed that also included children with SLI, some group differences (normal versus impaired) have been identified, but the nature of these differences warrants further study.

Understanding the ways in which a childhood language impairment manifests within the contrastive and noncontrastive forms of different dialects is a topic of ongoing study. Additional work also is needed to extend the study of childhood language impairment to other language-learning situations. Two such situations are bilingual language acquisition and second language learning. Until this research is completed, our tools for identifying children with language weaknesses and our understanding of language impairment as a construct will remain limited.

—Janna B. Oetting

References


Discourse

The term discourse is applied to language considerations beyond the boundaries of isolated sentences, although a discourse in its simplest form may be manifested as a single utterance in context, such as “Children at play.” Discourse studies emerge from a variety of disciplines, with major contributions from linguistics and psychology. Linguists have been motivated primarily by the desire to explain phenomena that cannot be accounted for at the word and sentence levels, such as reference or given/new information, while psychologists have emphasized strategic processes and the role of cognitive factors, such as memory, in the production and comprehension of discourse. This entry focuses on seminal linguistic research areas that have had a strong influence on the field.

A construct that defines the nature of discourse study, regardless of discipline or perspective, is coherence. A discourse is coherent when it “hangs together,” or makes sense. This notion of coherence pervades the approaches of a variety of discourse analysts. With some, it is used as a technical term in its own right. With others, it forms an underlying attribute of other constructs. Despite differences in terminology or focus, discourse approaches are similar in their analysis of an organization that supersedes any single sentence or utterance. Thus, the main goal of discourse analysis is to differentiate discourse from random sequences of sentences or utterances.

Discourse models generally assume that discourse coherence is realized through the integration of a variety of resources: the information contained in the text, shared knowledge, and the relevant features of the situation in which the text is embedded. Thus, linguistic form and meaning alone are insufficient for discourse comprehension and production. For that reason, discourse paradigms represent a dramatic shift from more traditional linguistic pursuits that focus exclusively on linguistic forms in isolation.

The knowledge structures that contribute to the formation of coherence are thought to be varied, conventional, and differentiated from each other by the kind of information they contain. For instance, story schema (Mandler, 1984) or superstructure (van Dijk and Kintsch, 1983) represent knowledge of the way in which events unfold in a story (narrative), and how a story typically begins and ends. Script knowledge (Schank and Abelson, 1977) specifies the sequence of steps in common everyday routines, such as going to a restaurant or making a sandwich. Knowledge structures also include knowledge of common patterns of conversational exchange, such as question-answer sequences (Schegloff, 1980).

Discourse analyses differ primarily in the degree to which they focus on the relative contributions of text, shared knowledge, and context. One influential model (van Dijk and Kintsch, 1983) is quite detailed in its account of the transformation of semantic content into cognitive information content. It represents the semantics of a text as a set of propositions. These propositions display coherence through inference at a local level, as microstructures, and at a global level, as macrostructures that represent the topic or gist of the text. The micro- and macrostructures constitute the text base, which is integrated with shared knowledge. The product of that integration is a representation of the events depicted in the text, conceptualized as the situation model. Thus, this is a model of discourse as a process involving transformation of information.

Linguists have also contributed greatly to our understanding of the various means by which discourse is coherently organized above the level of the sentence, and how the surface features of a text contribute to these higher levels of organization (Halliday, 1985). The research has centered largely on narrative, in part because of its universality as a genre and the fundamental nature of the temporal organization on which it is based. The extensive research on narrative has historical roots in Propp’s (1928/1968) analysis of folk tales (from the field of rhetoric) and Bartlett’s (1932) study of narrative remembering (from the field of psychology).

In linguistics, an early and influential framework for analyzing narrative structure was provided by Labov and Waletzky (1967) and extended in Labov’s later work (1972). The model is organized around the role of sentential grammar in discourse-level structure. In this framework, the verbal sequence of clauses is matched to the sequence of events that occurred, as the means by which past experience is recapitulated in the narrative. The overall structure of the narrative progresses from orientation to complicating action to resolution. An additional component of the overall structure is evaluation, which is expressed through a wide range of lexicogrammatical devices. Thus, this work established a fundamental approach to analyzing how an event sequence is realized in linguistic form. Contemporary adaptations of this seminal approach to narrative analysis are numerous (Bamberg, 1997).

Work on textual cohesion (Halliday and Hasan, 1976) provides another view of the relationship between discourse organization and its component linguistic units. Halliday and Hasan define cohesion as “the set of possibilities that exist in the language for making text hang together [as a larger unit]” (p. 18). Cohesive ties, which are surface features of text, provide the relevant semantic relation between pieces of the text. These ties can be lexical or grammatical and include devices such as reference, conjunction, and ellipsis. Thus, cohesion is simi-
lar to coherence in that it is a relational concept; text is cohesive when it is coherent with respect to itself. The notion of cohesion has been widely applied, especially in the area of reference.

Linguistic accounts of discourse genre are another means of analyzing how discourse can be coherent. Although narrative has been the most extensively studied genre, Longacre (1996) provides a typology of various discourse genres (e.g., procedural, expository, narrative) in terms of both underlying knowledge structure and their linguistic realization. In his framework, discourse is classified by the nature of the relationship between the events and doings in the discourse, the nature of reference to agents in the discourse, and whether the events happened in the past or are not yet realized. As is typical of linguistic approaches, Longacre further specifies the surface linguistic characteristics of each discourse type, such as the types of tense and aspect markings on the verb, the typical forms of personal reference, and the nature of the linkage between sentences for each discourse genre. For instance, he specifies how the setting of narrative usually contains stative verbs, while the complicating action contains action verbs. He further analyzes how the peak or climax of the story can be marked through devices such as shifts in tense or changes in length and syntactic complexity (Longacre, 1981).

An especially detailed account of how thinking is transformed into language is found in cognitive-linguistic work using narratives (Chafe, 1980). Chafe’s framework addresses ways in which the flow of thought is matched to units of language during the process of verbalization. Not only does he address lexicosyntactic contributions to discourse formation, he also considers the way in which syntax and intonation interact in discourse production.

Several of the discourse theorists explore the commonality and variability of narrative features across languages and cultures. These pursuits reflect an interest in universality that is pervasive in the field of linguistics in general. Contributions in this area include studies of cross-linguistic differences in expression of the basic features of narrative, such as verbs or the marking of reference (Longacre, 1996); ethnic linguistic devices used in the various narrative components, such as evaluation (Labov, 1972); variations in linguistic expression both within individuals and across individuals from different cultures (Chafe, 1980); and cultural presuppositions reflected in the content of narratives (Polanyi, 1989).

Conversational discourse represents a discourse type very different from the other discourse genres. Research on conversational discourse emphasizes the role of context and social interaction (e.g., Schiffrin, 1994). A basic unit of analysis for this genre is the speech act, a construct derived from early work in the philosophy of language (Austin, 1965; Searle, 1969). Speech acts are utterances defined by their pragmatic functions, such as making statements, asking questions, making promises, and giving orders. The sequence of speech acts can display a coherence that extends beyond any one speech act. On this basis, van Dijk (1981) proposes his notion of a macro-speech act, which consists of sequences of speech acts that function socially as one unit. The use of these speech acts during turn-taking in conversation is also rule-governed (Sacks, Schegloff, and Jefferson, 1974; Schegloff, 1982). Another important pragmatic framework that guides the overall coherence of information exchange is that of a cooperative principle (Grice, 1975). It subsumes the maxims of quantity, quality, relation, and manner, which guide the amount, clarity, and relevance of information required in conversations between interlocutors. Finally, some of the most recent and socially relevant work in conversational discourse extends the notion of context by focusing strongly on context in a variety of social settings. In this research, context is specified broadly to include participants and their relative roles in particular societal settings in a given culture (e.g., Tannen, 1994).

The advantage of discourse analysis lies in its potential to address both linguistic and cognitive factors underlying a range of normal and disordered communication performance.

See also discourse impairments.

—Hanna K. Ulatowska and Gloria Streit Olness

References


**Discourse Impairments**

The history of aphasiology dates back more than a century, but impairments of discourse abilities have only recently been described. Whereas changes in discourse abilities have always been part of the qualitative description of language following a brain lesion, the conceptual frameworks needed to identify impaired components of discourse in brain-damaged individuals have been available only since the late 1970s (Joanne and Brownell, 1990). Initial descriptions of discourse impairments essentially referred to traditional linguistic indicators, such as the noun-verb ratio or the percentage of subordinate clauses (e.g., Berko-Gleason et al., 1980; Obler and Albert, 1984). With increasing knowledge about the organization of the meaning conveyed in discourse, more specific descriptors have been introduced, such as coherence (e.g., Iriagray, 1973) or T-units (e.g., Ulatowska et al., 1983). However, those concepts and descriptors were not connected with a broader conceptual framework of discourse. Only recently have general integrative discourse models made it possible to link these various discourse components and to capture the different levels of cognitive processing needed in order to convey or understand verbal communication. This article summarizes discourse impairments associated with different pathological conditions with reference to these integrated frameworks.

**Levels of Discourse Processing**

Discourse is a set of utterances aimed at conveying a message among interlocutors. It can take many forms, such as narrative, argument, or conversation. Because it combines language components in a communicative context, discourse may be the most elaborate linguistic activity. The complexity of this activity can be captured through multilevel models, such as that proposed by
Carl Frederiksen (Frederiksen et al., 1990), in which each level of processing can be analyzed separately. Selective impairments of these levels, leading to distinct discourse patterns, may be used to differentiate among various adult populations with neurological disorders.

Seminal work by Kintsch and van Dijk (1978) largely inspired current integrated discourse models (for a review, see Mross, 1990). According to these models, discourse processing—production or comprehension—results from a number of cognitive operations that take place on four levels of representation:

- **Surface level**—Traditional linguistic units such as phonemes or graphemes, morphemes, words, and their combination into sentences constitute the surface level. Impairments at this level are described elsewhere in this book.
- **Semantic level**—Concepts expressed in discourse, along with the links between them, constitute the semantic level of processing. The smallest semantic unit is the microproposition, which is made up of a predicate (typically expressed by verbs or prepositions) and one or more arguments (typically expressed by nouns). Discourse meaning can thus be represented as a semantic network made up of a list of hierarchically related micropropositions. The main ideas of a discourse can be represented by macropropositions. On the receptive side, these macropropositions are constructed by applying rules in order to condense, eliminate, or generalize micropropositions. The latter are related through logical, inferential, or pragmatic links to the world depicted by the discourse. Mirror processing stages allow one to go from the main ideas of communicative intent to micropropositional discourse. The semantic level of discourse largely depends on the individual’s semantic memory (general knowledge) and is independent of the linguistic (surface) level.
- **Situational level**—The processing of micropropositions and the relations among them leads to the construction of a situation model based on the subject’s world knowledge. The situational level corresponds to the representation of the situation or events depicted in the discourse constructed by the interlocutor.
- **Structural level**—Finally, the structural level corresponds to the sequential and temporal organization of meaning units in a discourse. This level is known as the structure of a discourse and is identified as the discourse schema, script, or frame. It is at this level that distinctions among narrative, argumentative, procedural, or conversational discourse can be made.

**Impairments of Discourse by Level of Processing**

Surface-level impairments constitute the essence of aphasia as traditionally defined: phoneme, morpheme, word, and sentence impairments. The presence of impairments at this level usually makes it difficult to appreciate other levels of discourse. This explains why discourse impairments are most likely to be noticed in individuals without surface-level deficits, such as those with right hemisphere damage, traumatic brain injury, or early-stage Alzheimer’s disease.

Few studies have looked for impairments at the microstructural level. In one such study, Joanette et al. (1986) showed that both patients with right hemisphere damage and normal controls produced discourse with similar microstructure. Stemmer and Joanette (1998) confirmed this observation but found that individuals with left hemisphere damage tended to produce more fragmented micropropositions, lacking in arguments. This resulted in a disruption of the connective structure of discourse, which requires predicates and arguments to be connected in order to form a semantic network of propositions. Numerous other studies have looked at cohesion, which can be considered representative of the microstructural level. Cohesion refers to the quality of local relationships between the elements of discourse and is frequently expressed through linguistic markers such as pronouns and conjunctions. Patients with traumatic brain injury, dementia, and in some cases right hemisphere damage have been reported to produce incohesive discourse typically characterized by the use of vague words or pronouns without clear referents. The lack of cohesion prevents the interlocutor from knowing what the speaker is talking about. In an incohesive discourse, micropropositions are also incomplete, since arguments are neither present nor identifiable, thus leading to a break in local coherence. Because local coherence is not established, discourse with cohesive problems can be viewed as a disconnected or incomplete semantic network.

The relationships among the elements of discourse form the macrostructure of a discourse. At this level, a logical progression of ideas ensures the global coherence of the discourse. Several authors have reported impairments at this level in the narrative discourse of patients with right brain damage, dementia, and traumatic brain injury (Nicholas et al., 1985; Glosser, 1993; Coelho, Liles, and Duffy, 1994; Ehrlich, 1994). Frequently the problem lies in the absence of principal ideas. Such discourse is incomplete and difficult to interpret. Another problem occurs when discourse contains unexpected information, in which case it is referred to as tangential. In such cases the links among ideas are not explicit and do not seem logical; this may happen with right-hemisphere-damaged individuals. Often, individuals who produce tangential discourse do not stay with the topic and jump from one subject to another. Such behavior may be attributed to some pragmatic inability to take the interlocutor’s point of view into account and establish a common reference (Chantaraine, Joanette, and Ska, 1998).

Discourse processing also requires the elaboration of a situation model. The presence of such a model testifies to the fact that the individual is able to make inferences by bridging several pieces of information. This ability is frequently impaired in patients with diffuse lesions (traumatic brain injury and dementia) or right brain damage. In some cases, patients’ understanding is partial and remains at a superficial level. If such patients perceive contradictions, they often try to resolve them by
invoking a plausible explanation from their own experience rather than from the information included or presupposed in the discourse. The difficulty in integrating information into a situation model has been proposed as an explanation for the inability to understand jokes, metaphors, or indirect requests that has been reported in patients with right brain damage and traumatic brain injury. In the case of dementia, it is thought to result from degradation of the semantic system itself.

The structural level of discourse is the level at which information is organized with respect to a given script, such as a narrative, which has to contain minimally a setting, a complication, and a resolution. Preservation of the text structure can guide the production or comprehension of discourse. Although text structure is considered to be robust, patients with traumatic brain injury and dementia have shown impairments at this level. For example, individuals with Alzheimer’s disease may omit some components of the narrative schema, even when pictures are provided to support their production (Ska and Guénard, 1993). Script deficits in discourse are thought to result from impairments similar to those affecting routine activities of daily life through planning, organizing, selecting, and inhibiting information (Grafman, 1989). Patients with frontal lesions, for example, may exhibit difficulties in the sequential ordering of and hierarchical relations among actions belonging to a given script (Sirigu et al., 1995a, 1995b).

Discourse abilities and their impairments constitute a privileged component of communication and language that researchers can study in order to appreciate the interaction between so-called linguistic and other components of cognition. Although long addressed in purely descriptive terms, discourse impairments can now be understood with reference to comprehensive models of normal discourse processing. A description of discourse according to these models allows researchers to identify the levels characterized in impaired individuals with particular brain lesions. The relationship between the various levels of discourse impairment and the different brain diseases or lesions is gradually becoming clearer. The availability of such specific descriptions will lead to a better understanding of the communicative disability affecting individuals with discourse impairments and should help clinicians develop strategies in order to help affected individuals overcome this disability.

See also DISCOURSE.

—Bernadette Ska, Anh Duong, and Yves Joanette

References


Functional Brain Imaging

Several techniques are now available to study the functional anatomy of speech and language processing by measuring neurophysiological activity noninvasively. This entry reviews the four dominant methods, electroencephalography (EEG) and magnetoencephalography (MEG), which measure the extracranial electromagnetic field, and positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), which measure local changes in blood flow associated with active neurons. Each of these techniques has inherent strengths and weaknesses that must be taken into account when designing and interpreting experiments.

EEG and MEG respectively measure the electrical and magnetic field generated by large populations of synchronously active neurons with millisecond temporal resolution (Hamalainen et al., 1993; Nunez, 1995). Asynchronous activity cannot be easily detected because the signals produced by individual cells tend to cancel each other out rather than summing to produce a measurable signal at sensors or electrodes outside the head. The bulk of EEG and MEG signals appear to be generated not by action potentials but by postsynaptic potentials in the dendritic trees of pyramidal cells.

Although EEG has excellent temporal resolution, on the order of milliseconds, it is limited by poor spatial resolution because of the smearing of the potentials by the skull (Nunez, 1981). As a consequence, it is very difficult to identify the source of a signal from the distribution of electric potentials on the scalp. For any given surface distribution, there are many possible source distributions that might have produced the surface pattern—thus, the inverse problem has no unique solution. This complication is particularly significant where there are multiple generators, as is often the case in speech and language studies. The signals from different neural generators are mixed together in the potentials recorded at scalp electrodes.

EEG measures the electrical field produced by synchronous neural activity; MEG measures the magnetic fields associated with these electric current sources. There are important differences, however, between MEG and EEG signals. First, magnetic fields are unaffected by the tissue they pass through, so there is far less distortion of the signal between the source and the sensor in comparison to EEG (Hamalainen et al., 1993). Second, because most MEG is a measure of only the radial component of the magnetic field, MEG is effectively blind to activity that occurs in cortical areas that are oriented roughly parallel to the sensor (i.e., mostly gyral convexities). Conveniently for speech scientists, most of human auditory cortex is buried inside the Sylvian fissure, making MEG ideal for recording auditory or speech-evoked fields. MEG has a temporal resolution comparable to that of EEG. Theoretically, MEG has somewhat better spatial resolution than EEG because magnetic fields pass unaffected through the tissues of the head, but this benefit is partly cancelled by the greater distance imposed between MEG sensors and the brain. Source localization in MEG is still limited by the non-uniqueness of the inverse problem, which becomes increasingly troublesome as the number of signal generators increases.

Most EEG and MEG studies in speech and language use an event-related potential (ERP) design. In such a design, the onset of EEG recording is time-locked to the onset of an event—say, the presentation of a stimulus—and the resulting EEG response is recorded. Because the ERP signal is a small component of the overall EEG signal, the event of interest must be repeated several times (up to 100), and the responses averaged. Another increasingly popular use of electromagnetic responses involves mapping regional correlations (synchrony) in oscillatory activity during cognitive and perceptual processes (Singer, 1999), which has been suggested to reflect cross-region binding.

Unlike the electromagnetic recording techniques, hemodynamic techniques such as PET and fMRI measure neural activity only indirectly (Villringer, 2000). The basic phenomenon underlying these methods is that an increase in neural activity leads to an increase in metabolic demand for glucose and oxygen, which in turn appears to be fed by a localized increase in cerebral blood flow (CBF) to the active region. It is these hemodynamic reflections of the underlying neural activity that PET and fMRI measure, although in different ways.

PET measures regional CBF (rCBF) in a fairly straightforward manner (Cherry and Phelps, 1996): water (typically) is labeled with a radioactive tracer, oxygen 15; the radiolabeled material is introduced into the bloodstream, typically intravenously; metabolically active regions in the brain have an increased rate of blood delivery, and therefore receive a greater concentration of the radioactive tracer; the regional concentrations of the tracer in the brain can then be measured using a PET scanner, which detects the decay of the radioactive tracer. As the tracer material decays, positrons are emitted from the radioactive nucleus and collide with electrons. Such a collision results in annihilation of the positron and electron and the generation of two gamma rays that travel away from the site of the collision in opposite directions and exit the head. The PET scanner, which is composed of a ring of gamma ray sensors, detects the simultaneous arrival of two gamma rays on opposite sides of the sensor array, and from this information the location of the collision site can be determined. PET also measures other aspects of local energy metabolism using different labeled compounds, based on the same principle, namely, that the amount of the agent taken up is proportional to the local metabolic rate. Oxygen metabolism is measured with oxygen labeled with oxygen 15, and glucose metabolism is measured with a molecule similar to glucose called deoxyglucose labeled with fluorine 18. The spatial resolution of PET is ultimately limited by the average distance a positron travels before it collides with an electron, which is in the range of a few millimeters. In practice, however, a typical PET study has a spatial resolution of about 1 cm.
The temporal resolution of PET is poor, ranging from approximately 1 minute for oxygen-based experiments to 30 minutes for glucose-based studies.

Typical PET experiments contrast rCBF maps generated in two or more experimental conditions. For example, one might contrast the rCBF map produced by listening to speech sounds with that produced in a resting baseline scan with no auditory input. Subtracting the resting baseline map from the speech-sound activation map would yield a different map highlighting just those brain regions that show a relative increase in metabolic activity during speech perception. Many studies attempt to isolate subcomponents of a complex process by using a variety of clever control conditions rather than a resting baseline. Whereas this general approach has yielded important insights, it must be used cautiously because it makes several assumptions that may not hold true. One of these, the “pure insertion” assumption, is that cognitive operations are built largely of noninteracting stages, such that manipulating one stage will not affect processes occurring at another stage. This assumption has been seriously questioned, however (Sartori and Umilta, 2000). Another assumption of the subtraction method is that the component processes of interest have neural correlates that are to some extent modularly organized, and further that the modules are sufficiently spatially distinct to be detected using current methods. In some cases this assumption may be valid, but in others it may not be, so again, caution is warranted in interpreting results of subtraction-based designs. These issues arise in fMRI designs as well.

Experimental designs that do not rely on subtraction logic are becoming increasingly popular. Correlational studies, for example, typically scan the participants under several parametrically varied levels of a variable and look for rCBF patterns across scans that correlate with the manipulated variable. For example, one might look for brain regions that show systematic increases in rCBF as a function of increasing memory load or of increasing rate of stimulus presentation. Alternatively, it is possible in a between-subject design to look for correlations between rCBF and performance on a behavioral measure.

In order to increase signal-to-noise ratios in PET studies, data from several participants are averaged. To account for individual differences in brain anatomy, each participant’s PET scans are normalized to a standard stereotaxic space and spatially smoothed prior to averaging (Evans, Collins, and Holmes, 1996). Group averaged CBF maps are then overlaid onto normalized anatomical MR images for spatial localization. Group averaging does improve the signal-to-noise ratio, but it also has drawbacks. First, there is some loss of spatial resolution. This is important, not just in terms of localizing the precise site of an activation, but also in terms of the ability to detect activations in the first place: spatially smaller activations are less likely to be detected than larger ones, even if they are equally robust, simply because there is a reduced likelihood of small activations overlapping precisely in spatial location across subjects. A related drawback is that it is often hard to distinguish between a difference in activation level and a difference in spatial distribution.

fMRI also is sensitive to hemodynamic changes, but not in the same way as PET. fMRI is based on a rather surprising physiological fact: when a region of brain is activated, both CBF and the metabolic rate of oxygen increase, but the CBF increase is much larger. This means that the local venous blood is more oxygenated during activation, even though the metabolic rate of oxygen has increased. The physiological significance of this is still not understood, but one possibility is that the increased level of tissue oxygen is necessary to drive a higher oxygen flux into the mitochondria. The most commonly used fMRI technique is sensitive to these changes in the oxygen concentration of blood: this is the BOLD, or blood oxygenation level dependent, signal (Chen and Ogawa, 2000). The BOLD signal is intrinsic to the blood response, and so, unlike in PET, no radioactive tracers are needed. A typical fMRI experiment involves imaging the brain repeatedly, collecting a volume of images every few seconds for a period of several minutes, during which time the participant is presented with alternating blocks of two (or more) stimulus or task conditions. Brain areas that are differentially active in one condition versus another will show a modulation of the MR image intensity over time that correlates with the stimulus (task) cycles. Under ideal conditions the spatial resolution of fMRI comes close to the size of a cortical column, although in most applications the resolution is closer to 3–6 mm (Kim et al., 2000). The temporal resolution of fMRI is limited by the variability of the hemodynamic response. Under ideal conditions, fMRI appears to be capable of resolving stimulus onset asynchronies in the range of a few hundreds of milliseconds, and there is some indication that even better temporal resolution (tens of milliseconds) is possible (Bandettini, 2000). However, in most applications the temporal resolution ranges from about 1 s to tens of seconds, depending on the design.

Because fMRI measures intrinsic signals, it is possible to present the stimulus (task) conditions in short alternating blocks, or even as a series of individual events, within a single scan, unlike a PET study (Aguirre, 2000). Typical block design experiments present four to eight cycles of alternating blocks of different stimulus (task) conditions per scan. Event-related fMRI designs, which are modeled on ERP experiments, present stimuli individually rather than in blocks. This affords greater flexibility in experimental design: items consisting of different conditions can be randomly interleaved to decrease predictability of the upcoming items, and the blood responses to items can be sorted and averaged in a variety of ways, for example by accuracy or reaction time of simultaneously collected behavioral responses. The disadvantages of event-related designs include decreased amplitude of the response due to shorter stimulus durations, and increased sensitivity to regional differences in response onset, which can provide better information in the temporal domain but also can make it difficult to
model the hemodynamic response equally well across all activated brain regions.

fMRI is more sensitive than PET, allowing the detection of reliable signals in individual subjects. Despite this distinct advantage, most fMRI analyses are modeled on PET procedures, with spatial normalization of individual data sets, group averaging, and overlaying of activation maps onto normalized anatomical images. Also, as in PET experiments, most fMRI experiments utilize subtraction-based designs or correlational methods. Although fMRI has several advantages over PET, it also has several drawbacks related primarily to artifacts introduced into the signal from head motion, physiological noise (respiration, cardiac pulsations), and inhomogeneities in the magnetic field coming from a variety of sources. Another drawback, particularly relevant for speech/language studies, is the greater than 100 dB noise generated by the magnet during image acquisition. A potentially promising solution to this latter problem involves presenting auditory stimuli during silent periods between image acquisition (Hall et al., 1999).

—Gregory Hickok

References


Inclusion Models for Children with Developmental Disabilities

During 1998–99, 5,541,166 students with disabilities, or 8.75% of the school-age population ages 6–21 years, received special education and related services under Part B of the federal Individuals with Disabilities Education Act (IDEA) (U.S. Department of Education, 2000). IDEA specifies 13 disability categories based on etiological groupings. The largest single category of disability served is specific learning disabilities (50.8%), with speech and language impairments the second largest category (19.4%). Children with mental retardation account for 11.0%, while children with autism, considered a low incidence disability, constitute 1% of those receiving special education and related services. Since the original passage in 1975 of IDEA’s forerunner, the Education for All Handicapped Children Act, the categorical model has served as the basis for determining who qualifies for special education in accord with two premises. First, each disability category represents a separate and distinct condition that may co-occur with others but is not identical (Lyon, 1996), and second, dissimilar disability conditions, or deficits, require educational programs that differ from regular education programs.

The concept of inclusive schooling emerged during the 1980s and gained momentum in the 1990s in response to the two categorical premises of IDEA and its predecessor. Special education was viewed as a form of de facto segregation for children with disabilities. Furthermore, special education had assumed the mantle of a placement setting rather than seeing its primary function under federal regulations as the source of specially designed instruction and support services intended to meet children’s unique learning needs (Giangreco, 2001). Because of these reasons, major educational reforms were necessary to transform schools and instruction (Skr tic, 1992). Educational equity and excellence for all students required “a restructured system of education, one that eliminates categorical special needs programs by eliminating the historical distinction between general and special education” (Skr tic, Sailor, and Gee, 1996, p. 146). Inclusive schooling became the democratic mechanism proposed to accomplish this restructuring. From the perspectives of curriculum and instruction, inclusive practices in a single system of education should foster engaged learning. All classrooms should be learner-centered, children should be supported to become active and self-regulated learners, and instructional practices should be grounded to theme-based units, cooperative learning, team teaching, and student-teacher dialogue.
that scaffolds critical inquiry in both intellectual and social realms (National Research Council, 1999, 2000).

IDEA specifies that children with disabilities must be educated in the least restrictive environment. This means that, to the greatest possible extent, a child should be educated with children who do not have disabilities and that an explanation must be provided in the Individualized Educational Plan (IEP) of the degree, if any, to which a child will not participate in regular class activities (Office of Special Education and Rehabilitation Services [OSERS], 2000). In the practical implementation of this requirement in the past 10 years, the terms least restrictive environment, mainstreaming, and inclusion often have become confused (Osborne, 2002). Least restrictive environment is a legal requirement. It specifies that children can be removed from the regular education classroom for special education placement only if the nature and severity of the child’s disability are such that education in regular classes with the use of supplementary aids and services cannot be satisfactorily achieved (OSERS, 2000). In other words, maintenance in the regular education classroom means that children receive special education and related services for less than 21% of the school day. Mainstreaming is an educational practice that refers to the placement of students in regular education for part of the school day, such as physical education or science classes.

Inclusion is neither specified in law nor regulations as a placement. Similar to resource rooms and self-contained classrooms as optional placement settings, inclusion is also an option along the educational continuum of the least restrictive environment. In its broadest sense, inclusion is an educational philosophy about schooling. Inclusive schools are communities of learners “where everyone belongs, is accepted, supports, and is supported by his or her peers and other members of the school community while his or her educational needs are met” (Stainback, Stainback, and Jackson, 1992, p. 5). Full inclusion is the complete integration of the regular and special education systems where all children with disabilities receive their education, including special education and related services, as an integral part of the regular education curriculum. A major criticism of full inclusion models has been that placement decisions begin to take precedence over decisions about children’s individual educational needs (Bateman, 1995). In contrast, partial inclusion, which is more typical of current inclusive schooling, pertains to those situations in which one or more classrooms within a school or school district are inclusive. Most advocates of full inclusion would consider partial inclusion as inconsistent with the philosophy of inclusive schooling. Moreover, judicial tests of inclusion have primarily involved students with moderate to severe cognitive disabilities. In general, in case law situations, courts have ruled that “inclusionary placement” should be the placement of choice, with a segregated placement occurring only when the evidence is overwhelming, despite the school district’s best efforts, that inclusion is not feasible (Osborne, 2002).

Depending on state education laws, speech-language pathologists may provide either special education (instructional) services or, in conjunction with special education, related (support) services. The provision of related services must meet standards of educational necessity and educational relevance (Giangreco, 2001). In full and partial inclusion, special education and related services, as specified in an IEP, may be delivered through multiple models, all of which are premised on collaboration. In this sense, collaboration means that team participants have particular beliefs and possess certain skills (Silliman et al., 1999; Giangreco, 2000). These include (1) a shared belief in the philosophy of inclusion, (2) empowerment as decision makers combined with respect for varying decision-making values, (3) flexibility in problem solving about how best to meet the language and literacy needs of individual students, (4) the shared expertise made possible through coteaching strategies, and (5) high expectations for all students, regardless of their educational and disability status.

Based on the collaboration concept, the American Speech-Language-Hearing Association (ASHA, 1996) supported “inclusive practices” as an option for optically meeting the educational needs of children and youth. A flexible array of service delivery models for implementing inclusive practices was also specified in accord with children’s needs at different points in time. These four general models are not viewed as mutually exclusive or exhaustive. One model is direct service delivery in the form of “pull-out” services, considered appropriate only when there is a short-term objective for a child to achieve, such as acquisition of a new communication skill through direct teaching. A second model is classroom-based service delivery, in which the speech-language pathologist and teachers collaborate, for example through team teaching, to incorporate children’s IEP goals across the curriculum. The composition of team teaching models varies, but the teams may include a regular education teacher, a teacher of specific learning disabilities, and a speech-language pathologist who works full-time in the classroom (e.g., Silliman et al., 1999). A third model is community-based service delivery, in which communication goals are specifically addressed in community settings, such as a vocational education program that is a focus of a transition service plan. This plan is designed to bridge between the secondary education curriculum and adulthood and must be included in an IEP for students beginning at age 14 years (OSERS, 2000). The consultative model of service delivery is a fourth model; its main characteristic is indirect assistance. This model may be most appropriate when a child’s communication needs are so specific that they do not apply to other children in the classroom (ASHA, 1996).

For any inclusion model to be effective in meeting children’s diverse needs, two foundations must be functional. One concerns the change process that serves to translate conceptually sound research into everyday instructional practices. The process of changing beliefs and practices must be explicitly understood, supported,
and crafted to the particular educational situation (Skr tic, 1992; Gersten and Brengelman, 1996). The second fundamental support involves the building and sustaining of educational partnerships. A team’s capacity to sustain innovative and educationally relevant practices requires the successful integration of collaborative principles and practices (Giangreco et al., 2000).

Few studies have evaluated the outcomes of inclusion programs for students with language impairments. At least three predicaments have contributed to this situation. The first is the co-occurring disabilities dilemma. Few inclusion programs have been reported in the literature that specifically focus on children classified as having language impairment as the primary condition. The absence of data can be attributed in part to the categorical model, which fails to account adequately for co-occurring disabilities, much less the existence of category overlap. For example, the U.S. Department of Education (2000) continues to report that, for identification and assessment purposes, “learning disabilities and language disorders may be particularly hard to distinguish . . . because these two disabilities present in similar ways” (p. II-32). Procedural requirements contribute one part of this dilemma. State education agencies typically report unduplicated counts of children with disabilities. Only the primary disability category is provided. For example, specific learning disability, mental retardation, or autism is often reported as the primary condition, with language impairment classified as the secondary condition. However, when duplicated counts are available, such as the count of all disabilities for each child that the Florida Department of Education provides (U.S. Department of Education, 2000), language impairment emerges as the most frequent disability associated with another disability.

Second is the broad variations in research purposes and methods. Most outcome studies have primarily addressed the social benefits of inclusion for children with severe development disabilities, such as autism or Down syndrome. The results are complicated to evaluate because of significant disparities among studies in their definitions of inclusion, sample characteristics, such as ages, grades, gender, and type of disabilities, and the instructional or intervention focus (McGregor and Vogelsberg, 1999; Murawski and Swanson, 2001).

Two major issues confront the design of future research on the efficacy of inclusive practices (ASHA, 1996). First, in studying the cognitive and social complexities of teaching and learning, multiple research methodologies pursued in a systematic manner are necessary for exploring, developing, and testing hypotheses (Friel-Patti, Loeb, and Gillam, 2001) about the efficacy of the four inclusion models. Second, in moving past the social outcome focus, research strategies should examine individual differences in the ability to benefit academically and linguistically from inclusive education as a product of variations in instructional practices, expanding the scope of investigation beyond children’s performance on standardized measures of language and academic performance.

A third predicament is the broad variation in instructional practices in inclusive classrooms. The fact that disruptions in language and communication development are implicated in a wide range of disabilities but not acknowledged as central to children’s literacy learning (Catts et al., 1999) also has significant ramifications for research on academic outcomes of inclusion. For example, the effects of inclusion on emerging reading skill have been reported in a limited manner, primarily for children classified with learning (severe reading) disabilities (Klingner et al., 1998), less so for children with language impairment (Silliman et al., 2000). In general, the reading instruction for these children remained undifferentiated from the reading practices used with other students in the classroom, resulting in minimal gains. Thus, shifting a child to inclusion from a special education placement, or even maintaining a child in the regular education classroom, may mean that little has changed for that child in reality.

One implication for instructional practices is that educational team members, including speech-language pathologists, should be skilled in designing multilevel instruction that takes into account the individual child’s needs for the integration of oral language dimensions with evidence-based practices for learning to read, write, and spell (National Institute of Child Health and Human Development, 2000). A second implication, supported by ASHA (2001), is that speech-language pathologists are critical stakeholders in children’s literacy learning. They have the professional responsibility to bring their knowledge of language to the planning and implementation of prevention, identification, assessment, and intervention programs in order to facilitate children’s literacy development.

—Elaine R. Silliman

References


Bateman, B. D. (1995). Who, how, and where: Special education’s issues in perpetuity. In J. M. Kaufman and D. P. Hallahan (Eds.), The illus ion of full inclusion: A comprehensive critique of a current special education bandwagon (pp. 75–90). Austin, TX: Pro-Ed.


**Further Readings**


Many lines of evidence support the concept that the left hemisphere plays an essential and specialized role in language processing in adults. Studies of adults with focal brain injury find that approximately 95% of cases of aphasia are associated with left hemisphere damage (Goodglass, 1993). In the traditional view, damage to the third frontal convolution, Broca’s area, is associated with problems in language production, whereas damage to the first temporal convolution, Wernicke’s area, is associated with problems in language comprehension. However, the picture is more complex. The ability to predict the location of injury from the aphasic syndrome (or vice versa) is limited, and the right hemisphere remains involved in aspects of language processing, such as interpretation of prosody and metaphor and comprehension of complex syntax (Just et al., 1996). Nonetheless, the left hemisphere contribution to language seems to be necessary.

How, when, and why does the left hemisphere become specialized for language functions? The study of children who sustain focal left hemisphere damage prior to or during language development provides one experimental approach to address these issues. Rarely, children who have not yet learned to speak or who are still developing language skills sustain brain injury to areas of the left hemisphere that typically serve language function in adults. These children afford a naturalistic experimental opportunity to address the theoretical questions about the neural substrate of language learning.

If children with left hemisphere damage prior to language learning subsequently demonstrate serious delays in language development, the implication is that the mechanisms for language development reside within the damaged regions of the left hemisphere, a position called early specialization. Such findings would suggest that the neural architecture for language is determined by innate and probably genetic mechanisms. If, by contrast, children with left hemisphere damage successfully master language skills, the implication is that, at least under extreme circumstances, alternative organizations can be established. Such findings would suggest that the neural architecture of language is an outcome of language learning. In its strongest formulation, this second position asserts equipotentiality, that is, that either hemisphere can serve language functions as long as the neural commitment occurs before language learning forces left hemisphere specialization (Lenneberg, 1967). If children with left hemisphere damage show only minor delays, the implication is that alternative neural organizations are less favorable to language development or processing than the classical language areas (Satz, Strauss, and Whitaker, 1990), an intermediate position called constrained plasticity or ontogenetic specialization. This last position would suggest that some aspects of brain structure may be determined by early and possibly genetic factors, but that the full development of left hemisphere specialization emerges through development.

Challenges in the study of children with early focal brain injury complicate obtaining and interpreting findings. Before modern neural imaging modalities became available, the localization of damage was often uncertain; this review considers empirical studies that used computed tomography or magnetic resonance imaging for lesion location. Because early focal brain injuries are rare, published studies usually consist of small, heterogeneous samples. Age at onset, extent of lesion, time since injury, and associated problems vary within and across these groups, making meta-analyses impossible. The language capabilities of the infants, had they not sustained injury, remain uncertain. Other variables, including presence of seizure disorder or use of anticonvulsant medications, mediate outcomes (Vargha-Khadem et al., 1992). Given all the sources of potential variability, the use of group means to describe brain-injured populations may combine disparate groups and mask important distinctions. A preferable approach is individual profiling in the setting of a contrastive group of age- and developmentally matched children (Bishop, 1983).

An initial question to be addressed in considering the language development of children with focal lesions is their overall intelligence and cognitive profiles. If these children have severe intellectual impairments, then any language deficits should be evaluated in relation to cognitive measures such as intelligence quotient (IQ) tests. Most studies concur that children with focal injury to either hemisphere, even those with total surgical removal of a hemisphere, score at or near the population mean (Bates, Vicari, and Trauner, 1999). In children, unlike in adults, differential hemispheric mediation of verbal and performance or nonverbal functions is not typically found (Vargha-Khadem et al., 1992; Vargha-Khadem, Isaacs, and Muter, 1994; Bates, Vicari, and Trauner, 1999). The near-normal intellectual performance of
children with focal damage is a testament to the plasticity of the human brain.

**Development of Functional Communication Skills**

Children with left hemisphere damage to the classical speech areas are not aphasic. Individual differences occur in fluency, intelligibility, frequency of initiation, and volume of output. However, in light of the chronic sensory and motor deficits that follow injury in these children and the severe disruption of language that follows similar injuries in adults, it is remarkable that their conversational language is normal or near normal (Bates et al., 1997).

Despite the favorable prognosis, children with focal injury to either hemisphere may experience developmental delays in the onset of babbling and communicative gestures (Marchman, Miller, and Bates, 1991), vocabulary development, and use of word combinations in parent-child conversations (Thal et al., 1991; Feldman et al., 1992). Once these children begin to acquire functional skills, they are comparable in their rate of developmental progress to each other and to children developing typically (Feldman et al., 1992). By age 4, children with focal left hemisphere damage can master even the complex morphosyntactic structures of Hebrew, at least using a criterion that multiple complex structures are present in the spoken output (Levy, Amir, and Shalev, 1994).

Location of injury is not related to rate of development in the manner predicted from studies of adults with brain injury. Children with right hemisphere damage initially show greater delays in initial word comprehension and production than children with left hemisphere damage, and children with posterior left hemisphere damage, presumably including Wernicke’s area, develop more slowly than children with damage to other areas of the left hemisphere, including Broca’s area (Thal et al., 1991; Bates et al., 1997). These findings implicate the centrality of pattern recognition and subsequently of language comprehension in the development of language production.

Children with focal damage achieve high-level language skills. By school age, they can produce narrative discourse. Their productions are shorter and syntactically less sophisticated than those of age-matched peers, but narrative skills are comparable in children with left hemisphere and right hemisphere damage (Reilly, Bates, and Marchman, 1998). Most children with early focal damage also learn to read, write, and spell, although the strategies they use may vary as a function of the side of the lesion (Dennis, Lovett, and Wiegel-Crump, 1981).

These findings suggest that a wide neural network involving both cerebral hemispheres is necessary to launch language development, so that damage to a neural substrate in either hemisphere may delay language development. Once language skills begin to develop and an initial neural network is beginning to be established, neural organization progresses at a similar rate as in an intact system. The system is capable of high-level skills, such as narrative discourse and reading, but damage anywhere in the system may reduce the levels of functioning in these areas, presumably because these skills, even in maturity, require an extensive neural network.

**Performance on Specific and Formal Measures**

Despite adequacy of conversational language, children with left hemisphere damage show subtle to moderate impairments in select aspects of language. In children with injuries acquired during childhood, expressive and receptive complex syntax is particularly vulnerable (Aram, Ekelman, and Whitaker, 1986, 1987; Aram and Ekelman, 1987). At school age, an on-line sentence comprehension task suggested that their strategies for interpreting syntactic structures were developmentally delayed compared with those of normal learners (Feldman, MacWhinney, and Sacco, 2002). However, children with right hemisphere damage also showed developmental delays on this task. An alternative explanation to the interpretation that syntax skills may be particularly vulnerable to left hemisphere injuries is that subjects with left hemisphere damage have greatest difficulty with the most developmentally advanced areas on an experimental assessment (Bishop, 1983). In this regard, children who sustain left hemisphere damage in the perinatal period have also been shown to have more difficulties in lexical retrieval (Aram, Ekelman, and Whitaker, 1987), comprehension of difficult verbal material, and formulating sentences than children with right hemisphere injury (MacWhinney et al., 2000). They also show relatively more difficulties with reading, writing, and spelling than children with right hemisphere damage (Woods and Carey, 1978). The procedures that demonstrate the selective weaknesses of children with left hemisphere damage require precise and constrained performance, as opposed to the relatively free form of conversation.

On-line reaction time methodology has been used in school-age children with focal lesions to determine whether particular information processing abilities are selectively compromised in children with left hemisphere damage (MacWhinney et al., 2000). Children with both left and right hemisphere damage had slower reaction times than age-matched normal peers on all of the auditory and visual reaction time tasks studied. The two tasks that best distinguished children with left hemisphere damage from children with other lesions and normal children were verbally repeating numbers presented in the auditory mode and naming numbers presented in the visual mode, both tasks that require rapid verbal output.

**Age at Injury**

In most studies, the younger the child at the time of focal injury, the higher is the subsequent level of functioning in IQ and language. In children who sustain injuries after age 5, verbal IQ is more likely to be affected with left hemisphere damage and performance IQ with right
hemisphere damage. Children who undergo total hemispherectomy for intractable seizures have better language if the operation is done before 1 year of age than after 1 (Dennis and Whitaker, 1976). Nonetheless, Vargha-Khadem and colleagues (1997) reported the case of a previously nonverbal child who began to speak at age 9 years, after he had undergone a hemispherectomy of the left hemisphere for seizures and a reduction in anticonvulsant medications. The ability to develop language appears to be preserved under some circumstances into middle childhood.

**Brain Reorganization After Early Injury**

Language functioning can be relocated to the right hemisphere after early damage to the left hemisphere. Previously, the method used to determine the eloquent hemisphere was a sodium amytal carotid infusion, used in the presurgical evaluation of individuals with intractable seizures. If this anesthetic disrupts language when infused into the carotid artery on one side but not the other, then that side is considered the “eloquent” hemisphere. Rasmussen and Milner (1977) found evidence that individuals with early left hemisphere injury were far more likely to have language in the right hemisphere than individuals with no previous left hemisphere injury, although some individuals with early left hemisphere injury retained language functioning in the left hemisphere.

Functional imaging offers a noninvasive method to reexplore this issue. The methods are appropriate for normal individuals as well as clinical populations and can be used to determine the areas of activation for many different tasks. In adults, functional magnetic resonance imaging (fMRI) has shown that a wide network of areas is involved in sentence interpretation (Just et al., 1996); activation was more likely to include right hemisphere locations as the sentence difficulty increased. Booth and colleagues (2000) used a similar fMRI paradigm to compare six children with perinatal injuries, four with damage to left hemisphere areas, to normal adults and normal children. In adults and normal children, a sentence comprehension task produced more activation in the left hemisphere than in the right hemisphere; greatest activation in the superior temporal, middle temporal, inferior frontal, and prefrontal areas; and right hemisphere recruitment for difficult sentences, particularly in the adults. The children with left hemisphere damage performed less accurately than normal children on the task. These children activated primarily a right hemisphere network and did not show an increase in activation as a function of sentence difficulty. The children with left hemisphere damage also had very poor performance on a mental rotation task that typically activates right hemisphere areas. This finding suggests that reorganization of language to the right hemisphere may compromise skills typically served by the right hemisphere.

A variety of basic science methods show that cortical tissue can take on a variety of functions, suggesting that the cortex is pluripotential, if not equipotential. Synaptic connections seem to be formed and preserved on the basis of experience. Experience-dependent commitment of neural substrate may explain some of the variability in the neural organization of basic functions seen across individuals. Many studies also demonstrate experience-dependent progressive specialization of neural tissue. In language development, studies of normal infants and toddlers using event-related potentials have shown that in the initial development of a language skill, such as word learning or recognition of syntactic markers, a bilateral network becomes activated. As skill level increases, the function lateralizes to the left hemisphere (Mills, Coffey-Corina, and Neville, 1993, 1997).

The language of children with focal injuries leads to similar views on the neural basis of language learning. The development of language initially seems to use a wide bilateral neural network, such that damage to either hemisphere delays the onset. For reasons that remain unclear, a slight advantage to the left hemisphere for language function results in progressive specialization of the left hemisphere as learning proceeds. If the left hemisphere is damaged, then the other cortical areas specialize to serve language function.

Language functions may be preserved in adjacent regions of the left hemisphere or in homologous regions of the right hemisphere, depending on multiple factors. Given the pluripotential nature of cortex, children with left hemisphere damage perform well in conversational language. However, they have selective difficulties when tasks require syntactic skills and other rapid, precise, or constrained linguistic processing. However, children with right hemisphere injury may also have subtle to moderate language disturbances. The plasticity of the brain for language function may come about at the expense of other neuropsychological functions, including visual-spatial processing. Functional imaging holds enormous promise for investigating the organization of language and other skills in intact individuals developing typically as well as in children with focal injuries.

—Heidi M. Feldman

**References**


Further Readings


Language Disorders in Adults: Subcortical Involvement

The first suggestion of a link between subcortical structures and language was made by Broadbent (1872), who proposed that words were generated as motor acts in the basal ganglia. Despite this suggestion, according to the classical anatomo-functional models of language organization proposed by Wernicke (1874) and Lichtheim (1885), subcortical brain lesions could only produce language deficits if they disrupted the white matter fibers that connect the various cortical language centres. Consequently, aphasia has traditionally been regarded as
a language disorder resulting from damage to the language areas of the dominant cerebral cortex. Since the late 1970s, however, this traditional view has been challenged by the findings of an increasing number of cliniconeuroradiological correlation studies that have documented the occurrence of adult language disorders in association with apparently subcortical vascular lesions. In particular, the introduction in recent decades of new neuroradiological methods for lesion localization \textit{in vivo}, including computed tomography in the 1970s and more recently magnetic resonance imaging, has led to an increasing number of reports in the literature of aphasia following apparently purely subcortical lesions. (For reviews of \textit{in vivo} correlation studies, see Alexander, 1989; Cappa and Vallar, 1992, and Murdoch, 1996.) Therefore, although the concept of subcortical aphasia remains controversial, recent years have seen a growing acceptance of a role for subcortical structures in language. Despite an abundance of theoretical models, however, the precise nature of that role remains elusive.

Subcortical structures most commonly purported to have a linguistic role include the basal ganglia, the thalamus, and the subcortical white matter pathways. Some evidence for a role for the cerebellum in language has also been reported (Leiner, Leiner, and Dow, 1993). The basal ganglia comprise the corpus striatum (including the caudate nucleus and the putamen and internal capsule), the globus pallidus, the subthalamic nucleus, and the substantia nigra. Although these nuclei are primarily involved in motor functions, the corpus striatum and globus pallidus have frequently been included in models of subcortical participation in language. In addition, several thalamic nuclei have also been implicated in language, in particular the ventral anterior nucleus, which has direct connections to the premotor cortex and indirect connections to the temporoparietal cortex via the pulvinar. The basal ganglia and the thalamus are linked to the cerebral cortex by way of a series of circuits referred to as the cortico-striato-pallido-thalamo-cortical loops. The majority of contemporary theories specify these loops as the neuroanatomical basis of subcortical participation in language.

Although there is general agreement that critical white matter pathways and the thalamus are involved in language, controversy and uncertainty surround the possible linguistic role of striatocapsular structures. Although in vivo correlation studies have documented beyond reasonable doubt that language impairments can occur in association with lesions confined to the striatocapsular region of the dominant hemisphere, considerable variability has been reported in the nature and degree of these language impairments, with no unitary striatocapsular aphasia being identified (Kennedy and Murdoch, 1993; Nadeau and Crosson, 1997). Varied impairments have been noted in spontaneous speech, confrontation naming, repetition, auditory comprehension, and reading comprehension. A number of authors have suggested that a difference exists between the type of aphasia associated with anterior striatocapsular lesions compared to posterior striatocapsular lesions. For example, Naeser et al. (1982) noted that patients with capsular-putaminal lesions extending into the anterior-superior white matter typically had good comprehension and slow but grammatical speech. In contrast, those with capsular-putaminal lesions including posteri-or white matter extension showed poor comprehension and fluent Wernicke's-type speech, while those with anterior-superior and posterior white matter involvement were globally aphasic. Further support for this anterior-posterior distinction was provided by Cappa et al. (1983) and Murdoch et al. (1986). Despite this apparent consensus, several other studies have questioned the accuracy and utility of the anterior-posterior dichotomy by describing a number of cases in which the patterns of language impairment could not be accounted for in terms of this anatomical distinction (Kennedy and Murdoch, 1993; Wallesch, 1985).

In contrast to the striatocapsular lesions, language disturbances following thalamic lesions present a more uniform clinical picture, and it is generally accepted that a typical thalamic aphasia can be characterized by the clinical presentation. Most commonly the aphasia resulting from thalamic injury is of a mixed transcortical presentation, sharing some features with both transcortical motor and transcortical sensory aphasia (Cappa and Vignolo, 1979; Murdoch, 1996). The features of thalamic aphasia most commonly reported include preserved repetition, variable but often relatively good auditory comprehension, a reduction in spontaneous speech output, a predominance of semantic paraphasic errors, and anoma. Lesions of the dominant anterolateral thalamus (including the ventral anterior, ventral lateral, and anterior nuclei) have been highlighted as the loci of aphasic deficits, given that infarctions in this region more consistently lead to aphasic disturbances than lesions involving the posterior parts of the thalamus (Cappa et al., 1986).

Attempts to explain the clinical manifestations of subcortical aphasia have culminated in the formulation of several theories of subcortical participation in language. These theories, largely developed on the basis of speech and language data collected from subjects who have sustained cerebrovascular accidents involving the thalamus or striatocapsular region, have been expressed as neuroanatomically based models. Two models of subcortical participation in language have been quite influential. The first of these, the response/release/semantic feedback model (Crosson, 1985), proposes a role for subcortical structures in regulating the release of preformulated language segments from the cerebral cortex. According to this model, the conceptual, word-finding, and syntactic processes that fall under the rubric of language formulation occur in the anterior cerebral cortex. The monitoring of anteriorly formulated language segments, as well as the semantic and phonological decoding of incoming language, occurs in the posterior temporoparietal cortex. Language segments are conveyed from the anterior language formulation center to the posterior language center via the thalamus prior to
release for motor programming. This operation allows the posterior semantic decoding centers to monitor the language segment for semantic accuracy. If an inaccuracy is detected, then the information required for correction is conveyed via the thalamus back to the anterior cortex. If the language segment is found to be accurate during monitoring, then it is released from a buffer in the anterior cortex for subsequent motor programming. In addition to subcortical structures participating in the preverbal semantic monitoring process, the model also specifies that the striatocapsular structures are involved in the release of the formulated language segment for motor programming. Specifically, it is suggested that this release occurs through the cortico-striato-pallido-thalamo-cortical loop in the following way. Once the language segments have been verified for semantic accuracy, the temporoparietal cortex releases the caudate nucleus from inhibition. The caudate nucleus then serves to weaken inhibitory pallidal regulation of thalamic excitatory outputs in the anterior language center, which in turn arouses the cortex to enable the generation of motor programs for semantically verified language segments. According to this model, Crosson (1985) hypothesized that subcortical lesions within the cortico-striato-pallido-thalamo-cortical loop would produce language deficits confined to the lexical-semantic level.

Crosson’s (1985) original conception of the response-release mechanism has since been revised and elaborated in terms of the neural substrates involved (Crosson, 1992a, 1992b). Although the actual response-release mechanism in the modified version resembles that in the original conception, the route for this release is altered. The formulation of a language segment causes frontal excitation of the caudate, which increases inhibition of specific fields within the globus pallidus; however, this level of inhibition alone is not sufficient to alter pallidal output to the thalamus. An increase in posterior language cortex excitation to the caudate, which occurs once a language segment has been semantically verified posteriorly, provides a boost to the inhibition of the pallidum. The pallidal summation of this anterior and posterior inhibitory input allows the release of the ventral anterior thalamus from inhibition by the globus pallidus, causing the thalamic excitation of the frontal language cortex required to trigger the release of the language segment for motor programming. Overall, the revised model provides an integrated account of how subcortical structures might influence language output through a neuroregulatory mechanism that is consistent with knowledge of cortical-subcortical neurotransmitter systems and structural features.

A second model of subcortical participation in language was proposed by Wallesch and Papagno (1988). This model, referred to as the lexical selection model, also proposes that subcortical structures participate in language processes via a cortico-striato-pallido-thalamo-cortical loop. Wallesch and Papagno (1988) postulated that the subcortical components of the loop constitute a frontal lobe system comprised of parallel modules with integrative and decision-making capabilities rather than the simple neuroregulatory function proposed in Crosson’s (1985) model. Specifically, the basal ganglia system and thalamus were hypothesized to process situational as well as goal-directed constraints and lexical information from the frontal cortex and posterior language area, and to subsequently participate in the process of determining the appropriate lexical item, from a range of cortically generated lexical alternatives, for verbal production. The most appropriate lexical alternative is then released by the thalamus for processing by the frontal cortex and programming for speech. Cortical processing of selected lexical alternatives is made possible by inhibitory influences of the globus pallidus on a thalamic gating mechanism. This most appropriate lexical alternative has an inhibitory effect on the thalamus, promoting closure of the thalamic gate, resulting in activation of the cerebral cortex and production of the desired response. Cortical processing of subordinate alternatives is suppressed as a consequence of pallidal disinhibition of the thalamus, and the inhibition of cortical activity.

Despite the apparent differences in the two models, they both ascribe an important role to the subcortical nuclei in language processing, especially at the lexical level of language organization. It is equally apparent that each of these models has a number of limitations and that no one model has achieved uniform acceptance. A major limitation of these models is that neither explains the considerable variability in clinical presentation of subcortical aphasia. According to Cappa (1997), a further problem is that the models suggest such extensive and widely distributed systems subserving lexical processing that specific predictions appear to be difficult to disprove on the basis of pathological evidence. Put more simply, these models do not lend themselves readily to empirical testing. Yet another limitation arises from the nature of the research on which these models are based. The available models of subcortical participation in language are largely based on the observation that certain contrasting deficits of language production arise in subjects with particular subcortical vascular lesions when tested on traditional tests of language function. These language measures were typically designed for taxonomic purposes regarding traditional cortical-based aphasia syndromes and may be inadequate for developing models of brain functioning (Caramazza, 1984). It has also been argued that language deficits associated with subcortical vascular lesions may actually be related to concomitant cortical dysfunction via various pathophysiological mechanisms. For instance, cortical infarction may not have been detected by neuroimaging. Also, subcortical lesions may result in diaschisis or the functional deactivation of distant related cortical structures (Metter et al., 1983). Further, language dysfunction following subcortical lesions may be related to decreases in cortical perfusion, causing widespread cortical damage that may or may not be detected by neuroimaging (Nadeau and Crosson, 1997). As yet, however, the relationship between the structural site and etiology of subcortical lesions, the extent of
cortical hypometabolism and hypoperfusion, and associated language function remains to be fully elucidated.

Further clarification of the role of subcortical structures in language is likely to come through the use of functional imaging techniques and neurophysiological methods such as electrical and magnetic evoked responses, as well as from the study of the language abilities of patients with circumscribed neurosurgical lesions involving subcortical structures (e.g., thalamotomy and pallidotomy). Functional imaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) enable brain images to be collected while the subject is performing various language production tasks (e.g., picture naming, generating nouns) or during language comprehension (e.g., listening to stories). These techniques therefore enable visualization of the brain regions involved in a language task, with a spatial resolution as low as a few millimeters. The use of fMRI in the future is therefore likely to further inform the debate as to the role of subcortical structures in language. A review of the extensive literature on PET studies indicates that some studies published since 1994 have demonstrated activation of the thalamus and basal ganglia during completion of language tasks such as picture naming (Price, Moore, et al., 1996) and word repetition (Price, Wise, et al., 1996).

In summary, although a role for the thalamus in language is generally accepted, some controversy still exists as to whether the structures of the striatocapsular region participate directly in language processing or play a role as supporting structures for language. Contemporary theories suggest that the role of subcortical structures in language is essentially neuroregulatory, relying on quantitative neuronal activity. Although these theories have a number of limitations, for the present they do serve as frameworks for generating experimental hypotheses which can then be tested in order to advance our understanding of subcortical brain mechanisms in language.

—Bruce E. Murdoch

References


Further Readings

Wallesch, C. W., Johannsen-Harbach, H., Bartels, C., and legitimate, systematic, and rule-governed (LSA, 1997).

Society of America (LSA), which asserts AAE to be ASHA’s position is consistent with that of the Linguistic

American children. In addressing this issue, the Ameri-
glish (AAE), the term preferred here, diﬀers suﬃciently

number of these children speak a form of English vari-

interest in language disorders among African-American

children arises from the recognition that a signiﬁcant

number of these children speak a form of English vari-

ously referred to as Black English, African-American

English, African-American Vernacular English and

Ebonics (see dialect speakers). African-American En-

glish (AAE), the term preferred here, diﬀers suﬃciently

from Standard American English (SAE) to adversely

aﬀect the educational and clinical treatment of African-

American children. In addressing this issue, the Ameri-
can Speech, Language, and Hearing Association (ASHA)
has taken the oﬃcial position that children should not be viewed as having a speech and language
problem because they speak AAE (ASHA, 1983).

ASHA’s position is consistent with that of the Linguistic Society of America (LSA), which asserts AAE to be
legitimate, systematic, and rule-governed (LSA, 1997).

Despite proclamations of this kind, child speakers of
AAE are overrepresented in special education classes, in part because of their linguistic background (Kuelen,
Weddington, and Debose, 1998).

An important factor contributing to this over-
representation is clinicians’ failure to diﬀerentiate legiti-
mate patterns of AAE from symptoms of a language
disorder. This failure results from an assessment process
that relies heavily on identifying deviations from SAE as
signs of impairment. Moreover, when these deviations
are the sole symptoms and no conﬁrming evidence exists
of concomitant disorders such as hearing impairment,
cognitive-intellectual deﬁcits, neurological impairment,
or psycho-emotional problems, the reliance on deviant
SAE patterns for diagnosis is even greater.

A case in point is speciﬁc language impairment (SLI),
a disorder presumably restricted to aberrant language
symptoms without a known cause (Watkins and Rice,
1994). Because the language symptoms of SLI can
appear similar to legitimate language patterns of AAE
(Seymour, Bland-Stewart, and Green, 1998), African-
American children are at risk for SLI misdiagnosis. This
kind of misdiagnosis epitomizes linguistic and cultural
bias in assessment, which has been a major issue of con-
cern to clinical professionals committed to equity and
fairness in testing. Although far from resolved, linguistic
bias in testing has been addressed by focusing on three
related areas: language acquisition milestones for AAE,
reduction of bias in assessment methods, and reduction
of bias in intervention strategies.

Language Acquisition Milestones for AAE. Much of
what is known about language acquisition in SAE
undoubtedly also applies to the AAE-speaking child.
However, it is not altogether clear whether speakers of
AAE and SAE follow parallel tracks in mastering their
respective adult systems. Acquisition data on AAE sug-

gest that the two are quite similar until approximately
the age of 3, at which point they diverge (Cole, 1980).
This claim rests largely on evidence that young children
from both language groups produce similar kinds of
developmental “errors.” However, these similarities
may not occur for the same reasons, since several early
developmental patterns also appear to match the adult
AAE system. For example, absent morphological inﬂec-
tions are common in the emerging language of AAE and
SAE as well as in adult AAE.

Whether these early patterns are a function of devel-
oment or are manifestations of the AAE system is an
important question. It may be that AAE development
and maturation are uniquely inﬂuenced by adult AAE
in ways unlikely for SAE. Some preliminary evidence
to support this position comes from the work of Wyatt
(1995), who showed that African-American preschoolers
followed the same adult AAE constraint conditions in
their optional use of zero copulas. No comparable anal-
ysis has been done on zero copulas at earlier periods or
for developmental SAE patterns, however.

Evidence of diﬀerences in acquisition becomes clearer
as children’s language systems mature and AAE patterns
become more evident (Washington and Craig, 1994).
Features that once appeared similar between AAE and SAE begin to disappear in SAE and may even increase in frequency in AAE, as with the zero copula after the age of 3 (Kovac, 1980). Between the preschool period and age 5, archetypical AAE features are observed in children across socioeconomic levels, but their density is greater among low socioeconomic classes and among males (Wyatt, 1995; Washington and Craig, 1998).

Although the descriptive accounts of early AAE have provided important information about the characteristics and pervasiveness of child AAE, still limited milestone data exist about age ranges at which language structures are mastered and the appropriate form those structures should take. In contrast, a rich source of acquisition data in SAE establishes when children of various ages acquire language milestones in ways consistent with their SAE peers. This disparity in milestone data between AAE and SAE requires a somewhat different assessment strategy in order to reduce bias.

Reduction of Bias in Assessment Methods. Of the several kinds of possible bias in language disorders (Taylor and Payne, 1983), perhaps the most intractable is linguistic and cultural bias associated with existing standardized tests. These tests are biased because they have not been specifically designed for and standardized on AAE. As a consequence, alternative and “nonstandardized” assessment methods have been recommended (Seymour and Miller-Jones, 1981; Leonard and Weiss, 1983; Stockman, 1996). These methods include language sampling analysis and criterion-referenced language probes, which are both common methods in the clinical process and typically complement norm-referenced testing. Their specific use with AAE-speaking children is important because they offer a less biased, richer, more dynamic and naturalistic source for analysis than is found in the more linguistically biased, relatively restrictive, and artificial context of standardized tests.

However, there are disadvantages with language sampling and language probes as well. They are time-intensive and possibly less reliable, and they too are limited by the inadequate normative descriptions of AAE. In an attempt to minimize the importance of specific AAE norms in the assessment process, several authors have proposed focusing alternative assessment methods on those language elements that are not specific to AAE features. Such a focus circumvents difficult questions about the status of patterns such as absent language elements. Stockman (1996) proposed the Minimal Competency Core (MCC), which is a criterion-referenced measure that represents the lowest end of a competency scale of obligatory language patterns that typically developing children should demonstrate, irrespective of their language backgrounds. Similarly, Craig and Washington (1994) advocated the avoidance of several AAE-specific features dominated by morphosyntax by focusing on complex sentence constructions common to both AAE and SAE. Also, Seymour, Bland-Stewart, and Green (1998) showed that language features that did not contrast between AAE and SAE were better predictors of language disorders than those that were contrastive.

Each of the above recommendations can be useful in identifying possible language disorders. However, to determine the nature of the problem requires a more in-depth and complete analysis of the child’s language, since language disorders are likely to extend beyond only language behaviors shared between AAE and SAE, or only in complex sentences. To ignore AAE features or any aspect of the child’s language in determining the nature of a problem could yield an incomplete and distorted profile. Therefore, it is necessary to examine the child’s productive capacity for a variety of targeted language structures that have been identified in a representative sample of language and that can be probed further under various linguistic and situational contexts (Seymour, 1986). With sufficient evidence about the nature of a child’s language problems, the foundation then exists for intervention.

Reduction of Bias in Intervention Strategies. Decisions about intervention strategies depend directly on evidence obtained about the nature of the child’s problem. For reasons stated earlier, this evidence can be more valid and less biased when alternative or no standardized testing methods are used. However, because these methods are time-consuming and require a multiple phase process, Seymour (1986) advocated a diagnostic-intervention model in which intervention is part of ongoing assessment. In this model, intervention is based on diagnostic hypotheses formulated from an initial and tentative diagnosis, and then tested by language probes. The process is one of repeatedly formulating hypotheses, testing them, and reformulating them, again and again, as needed.

The test-retest approach is recommended for AAE-speaking children largely because of the uncertainty about the nature of AAE. This uncertainty is less a factor in identifying a language disorder, since identification can be made without focusing on AAE features. However, when determining the nature of the child’s problem and treating those problems, AAE features should not be avoided if a complete and accurate account of the child’s language is the objective. Consider the kind of diagnostic information needed for an AAE-speaking child who fails to produce any copulas, unlike optional copula use by his AAE-speaking peers. At least two intervention strategies are possible: (1) to apply an SAE model by targeting copulas wherever they are obligatory in SAE, or (2) to follow an AAE model and target copulas in a manner consistent with optional use. Unfortunately, no matter how desirable the latter course of action might be, it is unlikely without greater knowledge of the linguistic conditions that determine optional use.

Consequently, a default to an SAE model in situations where clinical solutions for AAE are not readily apparent may be inevitable until AAE is more fully described and viewed as a complete grammar comprised
of systems (Green, 1995), rather than as simply a list of structures defined by their contrast with SAE. A system's account requires answers to some complex linguistic and social questions about African-American children's development and use of language in a context characterized by linguistic duality.

—Harry N. Seymour

References


Further Readings


The Latino population encompasses a diverse group of people who self-identify as descendants of individuals who came to the United States from a predominantly Spanish-speaking country. Over the past decade, the Latino population in the United States has increased four times faster than the general population (Guzmán, 2001). It is estimated that the size of the Latino population will represent one-quarter of the U.S. population, or approximately 81 million Latinos, by the year 2050. The large growth of the Latino population is largely attributable to its high fertility rate (National Center for Health Statistics, 1999); a large proportion of the population is in the childbearing years, and families tend to be larger. Children under the age of 15 account for 30.5% of the Latino population.

The Latino population is linguistically diverse with respect to dialects and languages spoken. The dialects spoken in the country of origin and subsequently brought to the United States evolved from the different regional dialects spoken by the original settlers, the languages spoken by the native peoples of the Americas, and the languages spoken by later immigrants. There are two major groups of Spanish dialects, radical and conservative (Guitart, 1996). The radical dialects are spoken primarily in the coastal areas of Spanish-speaking countries and the Caribbean, while the conservative dialects are spoken in the interior parts of the countries. The dialects vary in phonology, morphosyntax, semantics, and pragmatics, with the most drastic qualitative differences seen in phonology and lexicon. The differences in morphosyntax are more quantitative than qualitative. The specific dialects spoken by Latino children will be influenced by the dialects spoken in their community. Other factors influencing the dialect spoken include the degree of contact with Spanish and English speakers, whether the speaker is learning both languages simultaneously or sequentially, and the prestige attached to the various dialects with which the individual comes in contact (Poplack, 1978; Wolfram, Adger, and Christian, 1999).

Speaking Spanish is one of the major ties that bind the Latino population, and approximately 80% of the population reportedly speaks it. The vast majority of the Latino population consider themselves bilingual; a small percentage is monolingual in either English or Spanish. Twenty-eight percent of the Latino population report that they “do not speak English well” or speak it “not at all” (U.S. Census Bureau, 2000). The number of monolingual Spanish speakers and bilingual English-Spanish speakers reflects the fact that 35% of the Latino population is foreign born and that the majority of foreign-born Latinos entered the United States in the last three decades. Continuous immigration and growth of the Latino population, coupled with greater acceptability of linguistic diversity in the United States, might reverse the previous trend, in which immigrants lost their native language by the third generation (Veltman, 1988). The more likely trend is for a continuous growth of a Latino population that is bilingual.

Bilingualism is not a one-dimensional concept. Rather, bilingualism may be viewed as existing on several continua representing different language competencies in form, content, and use of the language (Valdes and Figueroa, 1994). Collectively, these individual competencies will dictate the child’s linguistic proficiency in a language. The degree to which proficiency is exhibited in any one language at a particular point in time is influenced by the situation, the topic, individuals, and context (Romaine, 1995; Zentella, 1997). A shift in topic or a shift in participants may result in a switching of the code. This type of code switching is a verbal skill that requires a large degree of linguistic competence. Code switches are also made by less proficient speakers as a way to compensate for insufficient knowledge of one language.

Latino children exhibit varying degrees of proficiency in both English and Spanish. Given the pervasiveness of English language media, the use of Spanish by the majority of Latino families, and the communities in which Latinos are raised, it is doubtful that many Latino children reach school age as true monolingual Spanish speakers. Some of the children may be considered to be sequential bilinguals because their major exposure prior to entering school was to Spanish and their linguistic skills in English are minimal. These children’s main exposure to English will come when they enter school. Impressionistically, many of these children are indistinguishable from monolingual Spanish speakers (e.g., Spanish-speaking children in Mexico). However, differences become apparent when their Spanish is compared with the language spoken by true monolingual Spanish speakers (Merino, 1992). Children who have been exposed to both languages at home and who tend to communicate in both languages, the so-called simultaneous bilinguals, show a wide range of linguistic skills in English and Spanish by the time they reach the school-age years. However, their exposure to and use of Spanish makes even the most English-fluent members of this group different from their monolingual peers. In environments that do not foster the development of the child’s first language, language attrition occurs. Some language patterns attributable to language attrition are similar to patterns seen in children with language disorders (e.g., gender errors) (Restrepo, 1998; Anderson, 1999).

Given the large linguistic variability in the population, differentiating between a language difference (expected community variation) and a language disorder (communication that deviates significantly from the norms of the community; Taylor and Payne, 1994) is not simple. Our existing literature on language development in Latino children focuses primarily on a limited number of grammatical structures used by monolingual Spanish-speaking children (Gutierrez-Clellen, 1998; Goldstein,
2000; Bedore and Leonard, 2001). However, most Latino children are either bilingual or in the process of becoming bilingual, and therefore existing normative data on language acquisition by monolingual children (e.g., Miller and Leadholm, 1992; Sebastián and Slobin, 1994) do not accurately represent the language development of the majority of Latino children.

Assessments are further complicated by the fact that most of the available assessment protocols assume a high degree of homogeneity of exposure to the content of test items and to the sociolinguistic aspects of the testing situation. The cross-cultural child socialization literature suggests that Latino children’s home routines are not always compatible with the content or the routines typically required in a language assessment (Iglesias and Quinn, 1997). Thus, poor performance on a particular assessment may reflect lack of experience rather than a child’s inability to learn (Peña and Quinn, 1997).

The growing number of Latinos and their overrepresentation in statistical categories that place children at higher risk for disabilities or developmental delays (Arcia et al., 1993; Annie E. Casey Foundation, 2000; Iglesias, 2002) make it imperative that our assessment protocols not only accommodate linguistic differences across groups but also takes into account children’s experiences. A variety of assessment protocols that take into consideration the languages and dialects spoken and the child’s experiences have been suggested (Erickson and Iglesias, 1986; Peña, Iglesias, and Lidz, 2001; Wyatt, 2002). These protocols suggest the judicious use of standardized tests, taking into consideration norming samples and possible situational test biases, and consideration of alternative nonstandardized assessment procedures such as ethnographic analyses, criterion-referenced assessments, and dynamic assessments. Further, consistent with IDEA regulations (Individuals with Disabilities Act Amendments of 1997) and ASHA’s position statement on the assessment of cultural-linguistic minority populations (American Speech-Language-Hearing Association, 1985), the assessments need to be provided and administered in the child’s native language. In many cases this will require the examiner to be bilingual in Spanish and English or will require the use of qualified interpreters (Kayser, 1998).

The interpretation of the assessment results must take into consideration the growing literature on language development in Latino populations (Goldstein, 2000; with recognition that performance may differ from the expected norm because of a language difference rather than a language disorder. Although most children will show normal development in one or both languages, some will demonstrate weaker than expected performance in one or both languages. The data obtained must be carefully examined in the context of the languages and the dialects to which the child has been exposed and the experiences the child brings to the testing situation.

Intervention, if warranted, will require a culturally competent approach to services delivery in which the families’ belief systems, including views on disability, are respected and intervention approaches are culturally and linguistically congruent with those of the children’s families (Lynch and Hanson, 1992; Maestas and Erickson, 1992; van Kleeck, 1994; Iglesias and Quinn, 1997). The language of intervention should be based on the children’s linguistic competencies, parents’ preference, and functionality, not on the clinician’s lack of proficiency in speaking the child’s language. The bilingual literature on typical and atypical language learners strongly supports the notion that intervention should be conducted in the child’s strongest and most environmentally functional language (Gutierrez-Clellen, 1999). In many cases, this will mean using Spanish as the language of intervention. The skills gained in the acquisition of the first language will facilitate acquisition of the second.

—Aquiles Iglesias

References


Further Readings


Language Disorders in School-Age Children: Aspects of Assessment

Approximately 7% of kindergarten children have a primary language disorder characterized by a significant delay in the comprehension or production of spoken language. Although these children have normal intelligence and hearing and are free of obvious neurological deficits such as cerebral palsy and severe emotional disturbances such as autism, their limitations in spoken language often persist throughout childhood, adolescence, and well into adulthood.

In addition, children with this disorder, often called specific language impairment, frequently demonstrate limitations in written language development during the school-age years, including problems in decoding print, comprehending text, spelling, and producing essays for school assignments. Problems in the social use of language, particularly during peer interactions, frequently affect these children as well.

**Phonology**

Most typically developing children have mastered the sound system of their native language by 7 or 8 years of age. In contrast, children with language disorders may demonstrate errors in the production of speech sounds during adolescence. Common errors include distortions of sibilants (s, z) and liquids (l, r), reduction of consonant clusters (kl, sp), imprecise articulation during rapid speech (Johnson et al., 1999), and difficulty articulating polysyllabic words (thermometer, rhinoceros; Catts, 1986). Phonological errors can impair intelligibility and make the child overly self-conscious. For these reasons, they should be addressed during language intervention. Standardized tests that can be used to identify phonological disorders include the Arizona Articulation Proficiency Scale and the Goldman-Fristoe Test of Articulation.

Phonological disorders in young children are sometimes predictive of later problems in learning to read, particularly when additional deficits in phonological awareness are present. Problems in phonological awareness—the ability to analyze and manipulate the sounds of the language—commonly occur in school-age children with language disorders and underlie their difficulties in learning to decode and to spell words (Catts, 1993; Lombardino et al., 1997). Unfortunately, difficulties in decoding text can seriously hamper a child’s reading comprehension, just as difficulties in spelling can hamper the writing process. It is therefore very important that deficits in phonological awareness be addressed during language intervention.

Phonological awareness can be evaluated with a variety of tasks. For example, sound deletion requires the child to delete a particular sound or syllable in a word (“Say tall without the t.” “Say haircut without the hair”), and sound segmenting requires that the number of sounds or syllables in a word be counted (“How many sounds do you hear in crab?” “How many syllables do you hear in elephant?”). Standardized tests of phonological awareness, such as the Comprehensive Test of Phonological Processing, can also be used to identify deficits in school-age children.

**Syntax and Morphology**

The conversational speech of school-age children with language disorders is often characterized by utterances that are shorter and simpler than those of their peers with typical language development. For example, by age 10 or 12 years, a typical child can produce discourse such as the following: “The other day, while I was waiting for bus 28, which goes to the mall, my friend Harry, who lives in Brighton, stopped by and wanted to play a game of checkers. Finally, after we had played about an hour, the bus arrived and we had to stop, but I was glad because he had already beaten me four games.” These two sentences, containing 32 and 28 words, respectively, and six verbs each (both finite and nonfinite), would be well beyond the syntactic competence of school-age children with language disorders.

To express the same content, the child with a language disorder might need to employ a series of eight or ten shorter utterances. Although these utterances might be free of grammatical errors, it is unlikely that they would contain the advanced, low-frequency syntactic structures used by the typical child, such as the adverbial clause containing an adjective clause (while I was waiting for bus 28, which goes to the mall), the elaborated subject (my friend Harry, who lives in Brighton), the perfect aspect (had played, had beaten), or the adverbial conjunct finally to link ideas across sentences. In addition, the child with a language disorder might exhibit numerous false starts, hesitations, and revisions—“maze behavior” (Loban, 1976)—and may struggle to call up the details of the situation, producing a laborious and confusing message.

Similar problems can be observed in written language when children are asked to produce narrative, persuasive, or expository texts for school assignments. School-age children with language disorders characteristically produce shorter texts with fewer details, poorer organization, and a greater number of grammatical and spelling errors than their age-matched peers (e.g., Gillam and Johnston, 1992; Snowling, Bishop, and Stothard, 2000). They may also show evidence of morphological difficulties in writing. For example, they may omit the plural and past tense markers (“They use their skate yesterday”), or they may fail to use past irregular verbs correctly in obligatory contexts (“He teach them to read”) even as late as 12 years of age (Windsor, Scott, and Street, 2000). Problems in the use of derivational morphemes—prefixes and suffixes such as non-, -tion, and -ment—also occur in the speech and writing of children with language disorders (Moats and Smith, 1992).

The best way to evaluate a child’s syntax and morphology is to analyze spoken and written language samples with the aid of computer programs such as Systematic Analysis of Language Transcripts (Miller and Chapman, 2000). Standardized language tests can
also be administered, such as the Clinical Evaluation of Language Fundamentals–3, the Oral and Written Language Scales, the Test of Language Development–Primary, and the Test of Language Development–Intermediate.

**Semantics**

Compared to their age-matched peers, school-age children with language disorders frequently demonstrate limitations in lexical knowledge, particularly in relation to words that express abstract (pride, courage), polysemous (deep, absorbing), or technical (equation, parabola) meanings (Wiig and Secord, 1998). Figurative expressions such as metaphors (The lawyer was a bulldozer questioning the witness), idioms (throw in the towel, read between the lines), and proverbs (Every cloud has a silver lining) also pose comprehension difficulties, along with slang expressions (grandma lane), sarcasm ("Your room is SO clean now!"), and humor (Q: "Which sport is the quietest?" A: "Bowling. You can hear a pin drop") (Nippold and Fey, 1983; Lutzer, 1988; Spector, 1990; Milosky, 1994). Word retrieval, the ability to call up words with speed and accuracy, is often impaired as well, particularly in relation to low-frequency (tambourine) or abstract (religions) words (German, 1994).

Deficiencies in semantic development can seriously limit a child’s spoken and written communication, affecting social development and academic progress. For example, difficulties in word retrieval and humor comprehension can prevent a child from freely engaging in telling jokes and riddles, a popular pastime among school-age children. Because teachers’ classroom talk and the textbooks used in schools frequently contain difficult words and expressions, children with semantic deficits often fail to understand much of what they hear and read. These effects are cyclical, because listening and reading themselves are major sources of language-learning input during the school-age years. As a result, children who are deficient in listening and reading will continue to fall farther behind their peers in language development as they grow older.

Semantic development in school-age children can be evaluated through the use of standardized tests such as the Peabody Picture Vocabulary Test–III, the Test of Word Knowledge, and the Test of Word Finding–2. Informal observation of a child’s use of words, phrases, and expressions in social and academic contexts can also be informative. Deficiencies in listening and reading comprehension can be identified using standardized tests such as the Clinical Evaluation of Language Fundamentals–3 and the Woodcock Reading Mastery Tests–Revised, and by inspecting the child’s scores on academic achievement tests.

**Pragmatics**

Problems in pragmatics—the social use of language—commonly occur in school-age children with language disorders. Because of their phonological, syntactic, morphological, and semantic deficits, children with language disorders may receive social penalties from their typically developing peers in the form of teasing, hurtful comments, and personal rejection. This type of negative feedback can cause a child to avoid social situations and the opportunities they present for language development. For example, through regular interactions with peers, children are able to observe others using complex language and can themselves practice using appropriate phonological patterns, syntactic structures, morphemes, words, and figurative expressions in varied contexts such as greeting others, having conversations, exchanging information, agreeing and disagreeing, and persuading others to do things. A variety of pragmatic difficulties have been reported in school-age children with language disorders. In relation to peer interactions, these difficulties include limitations in the ability to access ongoing play groups; to collaborate, persuade, and negotiate; to engage in extended conversations; and to deliver bad news tactfully (Bliss, 1992; Brinton et al., 1997; Brinton, Fuji, and Higbee, 1998; Brinton, Fuji, and McKee, 1998; Fuji, et al., 2001).

Pragmatic development can be evaluated through role-playing tasks and social skills rating scales. Although the specific sources of pragmatic deficits are often unclear, in some cases they stem from limitations in the child’s social cognition—knowledge of the thoughts, feelings, and beliefs of others—and from deficits in the ability to use words, morphemes, and syntactic structures that mark politeness and empathy (e.g., “I don’t know why you weren’t chosen for the play, but I’m sure you’ll get the lead next time”) (Bliss, 1992). These problems, which can limit a child’s ability to maintain friendships, should be addressed in concert with other language goals.

See also LANGUAGE DISORDERS IN SCHOOL-AGE CHILDREN: OVERVIEW.

—Marilyn A. Nippold

**References**


Further Readings


**Language Disorders in School-Age Children: Overview**

Even though the symptoms and severity of a child’s language disorder may change over time, language disorders tend to be chronic. Preschoolers who are identified with language disorders are at substantial risk for experiencing language disorders during the school years and are also at risk for the academic, social, and vocational difficulties often associated with language disorders. Like younger children with language disorders, school-age children with language disorders are characterized by their heterogeneity. This heterogeneity manifests itself in the severity of the disorder, with some children showing mild grammatical difficulties, others showing no syntactic knowledge, and still others having no expressive language. For children with severe language disorders, spoken language may present inordinate difficulties. In these instances, children may use augmentative forms of communication, such as graphic systems, manual signs, and electronic speech output devices, to facilitate language development or to serve as alternate forms of communication.

The heterogeneity of language disorders in school-age children is also evident in the particular aspects of language that are disordered, with some children, for example, showing word-finding deficits, others having difficulty understanding complex directions, and yet others exhibiting global language deficits. Conventionally, a distinction is drawn between language disorders that affect only the production of language (expressive language disorders) and those that affect language comprehension in addition to production (mixed receptive-expressive language disorders). Children with either expressive or mixed language disorders may have a concomitant speech disorder, reflecting difficulty with speech sound representation and/or production. Disorders of reading and writing also may accompany language disorders (see LANGUAGE IMPAIRMENT AND READING DISABILITY).

For some children, language is the only developmental area in which they experience obvious difficulty; these children are often identified as having specific language impairment (SLI) (see SPECIFIC LANGUAGE IMPAIRMENT IN
CHILDREN). Omission of grammatical markers may be the most salient language characteristic in SLI, but it is not the only language deficit that may hinder a child's academic performance. In other children the language disorder is secondary to other cognitive, motor, or sensory disorders.

Several populations of school-age children are at risk for language disorders. These populations include children with developmental disabilities, such as children with MENTAL RETARDATION, AUTISM, or a pervasive developmental disorder, and also children in whom only subtle cognitive deficits are implicated. Among the latter are children with learning disabilities or disorders as well as children with attention deficit disorder, characterized by frequent instances of inattention and impulsiveness, and children with disruptive behavior disorder, marked by aggressive behavior or the violation of social norms. Children with hearing impairments are also at risk for language disorders. Although most school-age language disorders are developmental, children may have acquired language disorders resulting from closed head injuries, seizure disorders, or focal lesions such as stroke or tumors. Taken together, children with language disorders constitute a large group of students for whom language poses substantial difficulties.

About 5% of students in the United States show a learning disorder (American Psychiatric Association, 1994). Learning disorders are identified as disorders of reading, written expression, and mathematics. However, many children with learning disorders appear to have an associated difficulty with spoken language that substantially affects their ability to meet classroom language demands. The comorbidity of language disorders, learning disorders, and also attention deficit and disruptive behavior disorders is well-established. The overlap between disorders is at some level intuitive. For instance, children with attention deficit disorder often show deficits in executive functions, such as difficulties in goal setting, monitoring behavior, and self-awareness (Ylvisaker and DeBonis, 2000). These characteristics may have deleterious effects on the child's ability to deal with the complex language tasks encountered in the classroom.

Traditionally, a distinction was made between language delay and language deviance. For example, children with mental retardation were considered to show a delayed profile of language development, consistent with delay in other cognitive abilities. Children with autism were considered to show deviant language characterized by patterns not found for typically achieving children. Current research suggests that this global distinction does not fully capture the language profiles of children with language disorders. Contrary to the idea of simple delay, for example, children and adolescents with Down syndrome show greater deficits in expressive than in receptive language (Chapman et al., 1998). And contrary to the notion of overall language deviance, children and adolescents with autism have been found to produce narratives similar to those of children with mental retardation (Tager-Flusberg and Sullivan, 1995).

Across populations of children, difficulties in all domains of language, semantic, syntactic, and pragmatic, have been found. Current thinking in speech-language pathology, however, is not to address individual skills in isolation but to focus on broader aspects of the child's language and the learning environment that will best promote the child's current and future communicative success (Fey, Catts, and Larrivee, 1995). This includes recognizing the link between language, especially phonological awareness (awareness of the sound structure of words), and literacy skills (Catts and Kamhi, 1999). Oral narrative production is another area that has received attention, in part because the ability to tell a cohesive story rests on other language and cognitive skills and in part because good narrative skills seem to be associated with good academic performance (Hughes, McGillivray, and Schmidek, 1997). Also, children with language disorders are at risk for fewer and less effective social interactions than other children of the same age. Thus, the language foundations for social interaction, particularly conversational skills, constitute a major area to be addressed.

The school years cover a broad developmental range, and language disorders during adolescence are as important to identify as disorders occurring at earlier ages. However, language development is more gradual and individual in adolescence than it is in younger children, and identification of a disorder may be particularly challenging. Later language developments, such as the acquisition of figurative language (e.g., metaphors and idioms), advanced lexical and syntactic skills (e.g., defining abstract words, using complex sentences), analogical reasoning, and effective conversational skills, such as negotiation and persuasion, each develop over an extended period. At the same time, competence with these language skills is fundamental for dealing effectively with the academic and social curricula of high school. Adolescents with language disorders are at risk for dropping out of school or in other ways not making a successful transition to employment or university after high school. Thus, emphasizing adolescents' functional competence in social communication has been increasingly advocated.

Both standardized and nonstandardized measures are used to assess school-age language disorders in children and adolescents. Although below-average performance on standardized tests compared with chronological or developmental norms remains the primary way of identifying the presence of a language disorder, criterion-referenced assessments provide a more direct guide to intervention. In criterion-referenced assessments, the emphasis is on how well the child reaches certain levels of achievement rather than on how the child's language performance compares with that of other children of the same age. For instance, criterion-referenced assessments can be used to determine how well a child understands vocabulary used in classroom textbooks or how effectively a child initiates conversations with other children. (Paul, 2001, describes many standardized and criterion-referenced language assessments.)
Another form of nonstandardized assessment focuses on the underlying cognitive processing skills that potentially are linked to some language disorders. This evaluation includes tasks assessing verbal working memory (e.g., recalling an increasing number of real words), phonological working memory (e.g., imitating nonsense words), and auditory perception (e.g., discriminating speech and non-speech sounds). School-age children with language disorders have been distinguished from their age peers by lower accuracy on verbal working memory and nonword repetition tasks (Ellis Weismer, Evans, and Hesketh, 1999; Ellis Weismer et al., 2000). Dynamic assessment also has been advocated as an effective nonstandardized assessment strategy (Olswang, Bain, and Johnson, 1992). In dynamic assessment, aspects of a language task are altered systematically to examine the conditions under which a child can achieve optimal success. Thus, dynamic assessment can be used to determine a child’s potential for benefiting from intervention, and also what guidance or structure will be most helpful in intervention. Criterion-referenced, processing-dependent, and dynamic assessments may be especially important for children from culturally and linguistically diverse backgrounds, for whom many current standardized language tests may be inadequate or inappropriate (Restrepo, 1998; Craig and Washington, 2000).

See also SPEECH DISORDERS IN CHILDREN: SPEECH-LANGUAGE APPROACHES; SPECIFIC LANGUAGE IMPAIRMENT IN CHILDREN.

—Jennifer Windsor

References


Further Readings


McCauley, R. J. (1996). Familiar strangers: Criterion-referenced measures in communication disorders. Lan-
Language Impairment and Reading Disability

Reading is a language-based activity. As such, there is frequently an overlap between developmental language impairments and reading disabilities. Evidence of a relationship between these developmental disabilities has come from two perspectives. One has been the study of the language development of children with reading disabilities, and the other has been the investigation of the reading outcomes of children with spoken language impairments.

Language Problems in Children with Reading Disabilities

Numerous studies have documented that children with reading disabilities have problems in language development. Most investigations have involved concurrent examination of language problems in children with existing reading disabilities (Vogel, 1974; Bradley and Bryant, 1983; McArthur et al., 2000), while a few have studied the early language abilities of children who later became reading disabled (Scarborough, 1990; Catts et al., 1999). The latter approach is critical to determining the direction of causality.

Children who become poor readers often have language problems, or at least a history of these problems. Poor readers may have difficulties in vocabulary, grammar, or text-level processing (Vogel, 1974; Catts et al., 1999; McArthur et al., 2000). In at least some cases, these deficits are severe enough for children to have been identified as language impaired (Catts et al., 1999; McArthur et al., 2000).

In addition to these language deficits, children with reading disabilities have difficulties in other areas of language processing, specifically phonological processing (Bradley and Bryant, 1983; Fletcher et al., 1994; Catts et al., 1999). The most noteworthy of these deficits are problems in phonological awareness. Phonological awareness is the explicit awareness of, or sensitivity, to the sounds of speech. It is one's ability to attend to, reflect on, or manipulate phonemes. Children with reading disabilities are consistently more impaired in phonological awareness than in any other single ability (Torgesen, 1996). Poor readers often have difficulties making judgments about the sounds in words or in their ability to segment or blend phonemes. Such problems make it difficult for children to learn how the alphabet represents speech and how this knowledge can be used to decode printed words.

Children with reading disabilities have also been reported to have other deficits in phonological processing (Wagner and Torgesen, 1987). Poor readers have problems in phonological retrieval (i.e., rapid naming), phonological memory, and phonological production (reviewed in Catts and Kamhi, 1999). Although these difficulties in phonological processing often co-occur with those in phonological awareness, there are notable exceptions. For example, one current theory proposes that phonological awareness and rapid naming are somewhat independent, so that poor readers may have deficits in either area alone or in combination (Wolf and Bowers, 1999). However, it is proposed that children with deficits in both areas, or what is termed a double deficit, are at greatest risk for reading difficulties.

Although most children with reading disabilities have a history of language problems, the overlap between language and reading disabilities is not complete. In each group of poor readers who have been studied, at least some participants do not appear to have a history of language problems. When language problems are defined on the basis of difficulties in vocabulary, grammar, or text-level processing, about half of poor readers show no evidence of a language impairment (Mattis, 1978; Catts et al., 1999; McArthur et al., 2000). However, when phonological processing deficits are also included, the percentage of unaffected poor readers is about 25%–30% (Catts et al., 1999). These results are due, at least in part, to the fact that other nonlinguistic factors likely contribute to reading disabilities. Current theories include visual deficits and speed of processing problems as alternative or additional causes of reading problems (Eden et al., 1995; Nicholson and Fawcett, 1999). However, the lack of an apparent association between reading disabilities and language impairments may also be the result of discontinuities in the growth of various aspects of language and reading abilities. These discontinuities may obscure the relationship between

---


language and reading disabilities at certain points in time and highlight the association at others (Scarborough, 2001).

**Reading Outcomes in Children with Language Impairments**

The overlap between language impairments and reading disabilities has also been established by studies of the reading outcomes of children with language impairments. In the earliest of these studies, children with a clinical history of language impairments were located later in childhood or adulthood and their academic achievement was compared with their earlier speech-language abilities (Aram and Nation, 1980; Hall and Tomblin, 1978). More recently, studies have identified children with language impairment in preschool or kindergarten and followed them into the school grades (Bishop and Adams, 1990; Catts, 1993; Stothard et al., 1998; Rescorla, 2000; Catts et al., 2002). Both lines of research indicate that children with a history of language impairments are at high risk for reading problems. In almost every instance, the reading outcomes of children with language impairment have been found to differ significantly from those of children with typical language development. In addition, many children with a history of language impairments could be classified as reading disabled. Across studies, the percentage of children with language impairment who have been found to have subsequent reading problems has varied from approximately 40% to 90%, with a median value of about 50%.

Despite a strong tendency for children with language impairment to develop reading problems, not all of these children become poor readers. Research indicates that the type and severity of the language disorder are related to reading outcome. Children with more severe or broader-based language impairments are at greater risk for reading disabilities than those with less severe problems or problems confined to a single dimension of language (i.e., expressive language) (Rescorla, 2000; Catts et al., 2002). Also, children with language impairments who have concomitant nonverbal cognitive deficits (i.e., low nonverbal IQ) have poorer reading achievement than those with normal nonverbal cognitive abilities (Bishop and Adams, 1990; Catts et al., 2002). Phonological processing abilities are also related to reading outcome in children with language impairments. For example, Catts (1993) found that measures of phonological awareness and rapid naming were predictive of reading achievement, especially word recognition abilities, in children with language or articulation impairments.

The persistence of a language impairment may be an important predictor of reading outcome in many children. Some children with language impairment appear to resolve their language difficulties prior to school entry, while others continue to manifest language impairments into the school years. Those who continue to have significant language problems are at a much greater risk for reading disabilities in the early school grades than those with improved language abilities. Bishop and Adams (1990), for example, reported that children with language impairments at age 4 who were no longer language impaired at age 5½ had normal reading achievement at age 8. Catts et al. (2002) further found that children with language impairments in kindergarten who did not show language impairments in second grade had significantly better reading outcomes than those who continued to have language problems. Finally, the “persistence hypothesis” is also supported by evidence that change in language impairment status is related to severity of language difficulties, type of language impairment, and nonverbal IQ, each of which has been associated with reading outcome in children with language impairment (Bishop and Adams, 1990, Catts et al., 2002).

These conclusions, however, are compromised somewhat by long-term follow-up data. Specifically, Stothard et al. (1998) showed that the children studied by Bishop and Adams (1990) who had improved in language abilities by age 5½ and who did not have reading problems at age 8½ subsequently did have language and reading problems when tested at age 15. In fact, 52% were found to read significantly below grade level. These results are consistent with Scarborough’s proposal of “illusory recovery” (Scarborough, 2001). According to this proposal, children who appear to have resolved their language problems early on, later show language impairments and demonstrate significant disabilities. Scarborough (2001) has argued that illusory recovery and apparent relapse may be the result of nonlinear growth in language development. She suggests that different aspects of language are characterized by spurts and plateaus in growth. Thus, individual differences in language development may be more apparent at some stages than others.

Nonlinear growth in different aspects of reading development may also influence the observation of a relationship between language and reading disabilities. For example, in the early stages, reading development is characterized by rapid improvement in word recognition skills, which rest heavily on phonological processing abilities. At later stages, individual differences become more related to language comprehension abilities (Hoever and Gough, 1990). Thus, children with deficits in phonological processing will most likely have problems in the early stages of learning to read, while those with vocabulary and grammar deficits will be especially at risk in the later stages of reading development.

In summary, the two lines of research reviewed here converge in support of a close relationship between language impairments and reading disabilities. However, this relationship is not complete and may be complicated by nonlinearities in both language and reading development.

—Hugh W. Catts

**References**


Language Impairment in Children: Cross-Linguistic Studies

For many years the study of language impairment in children focused almost exclusively on children learning English. The past decade has seen a decided and salutary broadening of this exclusive focus as researchers from many nations have become interested in understanding the impaired language acquisition of children speaking diverse languages. The concern for understanding the cross-linguistic nature of developmental language impairment has followed in the path of investigations into the cross-linguistic nature of typical language development that were initiated by Slobin and his colleagues (e.g., Slobin, 1985).

Cross-linguistic studies of language impairment in children have both a theoretical and a practical impetus. Theoretically, the study of typical and impaired language development across a number of structurally different languages should reveal commonalities as well as differences in how children learn languages. Such studies should help researchers determine what is universal and what is variable in the ways that children learn languages. By sorting the variable from the universal, such research enhances knowledge about the properties of language development that are general to the process of language learning and those that are determined by the structure of the language that the child is exposed to. Practically, understanding what form impairment takes in various languages improves the possibilities for assessment and treatment in those languages. Understanding the contribution of input to the nature and timing of acquisition in a specific language has important consequences for developing intervention approaches. On the other hand, understanding the common properties of language learning helps to clarify in what ways impairment results from possible disruptions of basic human biological and cognitive mechanisms.

Most cross-linguistic studies of children with language impairment have been concerned with the morphological deficits experienced by children with specific language impairment (SLI) (see SPECIFIC LANGUAGE IMPAIRMENT IN CHILDREN). Some such grammatical deficits are found to occur in almost all languages that have been studied so far, regardless of their structure. Other deficits are particular to individual language families. One deficit that appears to exist in many languages has to do with the acquisition of verbal inflection. Finite verbal inflections and auxiliary verbs produced by children with SLI are optionally missing or substituted for in the following languages: German (Clahsen, Bartke, and Gollner, 1997; Rice, Ruff, Noll, and Grimm, 1997), Dutch (de Jong, 1999), Swedish (Hansson, 1997), Norwegian (Meyer Bjørkan, 1999), English (Rice and Wexler, 1996), French (Jakubowicz and Nash, 2001; Paradis and Crago, 2001), Italian (Bottari, Cipriani, and Chilosi, 1996; Bortolini, Caselli, and Leonard, 1997; Bottari et al., 2001), Japanese (Fukuda and Fukuda, 2001), Greek (Clahsen and...
Various language families show interesting patterns that relate to the structure or typology of those particular groups of languages. For example, studies of Germanic languages such as Dutch, Swedish, and German have shown that there are word order consequences related to the omission of finite verb inflections and auxiliaries in the speech of children with SLI.

Studies of Romance languages such as Italian and French have demonstrated that children with SLI have greater trouble acquiring the past tense than the present tense, while English-speaking children with SLI have difficulty with both tenses. Impaired speakers of French and Italian also have difficulty with the production of object clitics. Surprisingly, however, Italian and French children with SLI differ in their acquisition of determiners. Italian-speaking children have more difficulty with this aspect of their grammar than the French-speaking children. It is unfortunate that in the family of Romance languages, there are no comparable acquisition studies of children with SLI learning Spanish, either in the Americas or in Europe.

Speakers of non-Indo-European languages such as Inuktitut and Arabic have different patterns of impairment than other groups. For instance, Arabic speakers replace incorrect verbal inflections with a default form that is typically tense-bearing in the adult language. This is different from the Germanic and Romance languages, where tense-bearing morphemes, such as verbal inflections or auxiliary verbs, tend to be dropped, resulting in a nonfinite verb form such as an infinitive, a participle, or the verb stem appearing as the main verb in the sentence. Specific language impairment in Inuktitut has been shown to present yet another dimension. Here, the trouble a child with impairment had with verbal inflection did not resemble younger normally developing children. This is a different pattern than has been found in many other languages where children with SLI show a pattern of optional verbal inflection that resembles that of younger normally developing children.

Cross-linguistic studies can also play a particularly useful role in verifying hypotheses about the nature of the deficits experienced by children with SLI. They allow researchers to check out explanations based on particular languages with results from typologically different languages. For instance, Leonard, one of the pioneers of cross-linguistic studies of child language impairment, hypothesized from his study of English-speaking children that a plausible explanation for the grammatical deficit in SLI was that these children had difficulty establishing a learning paradigm for morphology in languages where the morphemes had low phonological saliency. He and his colleagues sought systematic cross-linguistic verification for this explanation by studying children with SLI who were learning languages with more salient morphology, such as Italian (Leonard et al., 1992) or Hebrew (Dromi, Leonard, and Shteiman, 1993; Dromi et al., 1999). Indeed, such children appeared to have less difficulty with the acquisition of their verbal morphology. However, succeeding studies of other languages with very phonologically salient verbal inflections, such as Inuktitut and Japanese, have shown that children with SLI did, in fact, have difficulty with the acquisition of verbal inflections. This demonstrates how a series of cross-linguistic studies can be useful in establishing whether a certain theoretical explanation of SLI is meaningful across languages.

Even though the number of languages that are being studied is expanding, there are only a few studies that can be considered truly cross-linguistic in design. These few genuine cross-linguistic studies involve either a format that compares in one study two groups of subjects who speak two different languages or a specific language in which the grammatical variables tested are as identical as possible to a previously studied language. In addition, studies in which investigators have examined the grammatical properties of children speaking a single specific language also display some methodological shortcomings impeding well-founded conclusions. They have varied in the age and criteria of selection of the children being studied. They have also varied in the particular properties of the children’s grammar that were assessed. Regrettably, such design issues have made the kinds of comparisons that are essential to establishing the universal and variable properties of acquisition and impairment often inconclusive.

In summary, despite certain limitations, the expansion of studies of language impairment in children to include a wider variety of languages has enlivened theoretical debate, led to new, linguistically based understandings, and enriched perspectives on this communicative disorder. It is important that studies of language impairment in childhood continue to encompass more and different languages. In fact, languages spoken by the vast majority of children in the world’s population remain virtually unexplored. It is equally important that more studies be designed to be truly cross-linguistic in nature, with variables and criteria for impairment established as uniformly as possible. The study of bilingual children with impairment provides a unique opportunity for the cross-linguistic observation of language impairment. These children are perfectly matched to themselves as the learners of two languages (see BILINGUALISM AND LANGUAGE IMPAIRMENT). Finally, just as important as understanding how children with language impairment learn different languages is recognizing that they are also members of different cultures. Language and culture are inextricably linked and as such they influence each other as well as the manifestations of and beliefs about childhood language impairment.

—Martha Crago and Johanne Paradis

References


Further Readings


Language in Children Who Stutter

A connection between language and stuttering in young children is intuitive. As noted by Yairi (1983) and others (e.g., Ratner, 1997), stuttering first appears in children between ages 2 and 4 years, during a time of rapid expansion in expressive and receptive language ability. Moreover, the repetitions and prolongations that characterize stuttering are observed as the child uses sounds to form words and words to form phrases and sentences. The apparent link between domains has given rise to theoretical accounts of stuttering that emphasize linguistic variables. For example, one working account of stuttering suggests that underlying difficulties with phonological encoding, difficulties that self-correct prior to actual language production but that slow language processing, yield disfluencies (Postma and Kolk, 1993). Linguistic factors are implicated in several other theoretical accounts of stuttering as well (Wingate, 1988; Perkins, Kent, and Curlee, 1991).

Despite the intuitive appeal of connections between language and stuttering, many of the most fundamental questions in this area of inquiry continue to be debated. The language abilities of young children who stutter have been the focus of research and controversy for many years (see Yairi et al., 2001, and Wingate, 2001, for examples of the ongoing dialogue on this topic). In addition to examinations of the language development status of young children who stutter, the connection between language and stuttering in young children has...
been studied in other ways, namely, through evaluation of linguistic variables that appear to exert an influence on stuttering behavior. Relevant linguistic variables include grammatical complexity and location of stuttering events in the language planning or production process. In both areas, the developmental status of children who stutter and the study of linguistic influences on stuttering, a growing body of knowledge speaks to the associations of language and stuttering in young children. This article highlights key research findings in these areas and summarizes what is known about the interface of language and stuttering in young children.

Language Ability and Stuttering in Young Children

The scholarly literature reveals a relatively longstanding view of the child who stutters as more likely to have language learning difficulties or impairments than typically developing peers. Through analysis of spontaneous language sample data, a group of scholars has empirically evaluated the expressive language abilities of a large cohort of young children who stutter (Watkins and Yairi, 1997; Watkins, Yairi, and Ambrose, 1999). The Illinois Stuttering Research Project has prospectively tracked a group of young children who stutter, beginning as near stuttering onset as possible and continuing longitudinally for a number of years to monitor persistence in versus recovery from stuttering. This work has focused on expressive language abilities, comparing the performance of young children who stutter with normative expectations on a range of language sample measures, such as mean length of utterance (MLU, a general index of grammatical ability), number of different words (NDW, a general measure of vocabulary skills), and Developmental Sentence Score (DSS, an index of grammatical skills). The researchers found that, as a group, children who stutter perform at or above normative expectations in their expressive language skills. More specifically, Watkins, Yairi, and Ambrose (1999) reported data based on analysis of 83 preschoolers who stuttered. Children who entered the study between the ages of 2 and 3 years (i.e., exhibited stuttering onset between ages 2 and 3 years) scored about 1 SD above normative expectations on several expressive language measures calculated from spontaneous samples. Children who entered the study between the ages of 3 and 4 years or 4 and 5 years performed at or near normative expectations. Interestingly, the children whose stuttering would ultimately persist (roughly 25% of the total sample) did not differ in expressive language skills from the children who would later recover from stuttering, when their language skills were compared near the time of stuttering onset. Figure 1 provides a sample of the findings of Watkins, Yairi, and Ambrose (1999), showing the MLU for children who entered the longitudinal study at three different age groupings relative to normative expectations.

The findings of several other investigators lend support to these results regarding expressive language skills in young children who stutter. More than a decade ago, Nippold (1990) reported that there was no compelling evidence that children who stuttered had a higher rate of language learning difficulties than the general population. In 2001, Miles and Ratner reported the performance of a group of young children who stuttered on a range of expressive and receptive language measures; the children in their sample scored at or slightly above the average level of performance on every reported measure (i.e., the group scored at or above a percentile rank of 50 or at or above a standard score of 100 on the measures used). Rommel and colleagues (1999) in Germany reported that preschool-age participants who stuttered had language skills at or above age expectations. Anderson and Conture (2000) also reported language scores at or above average for a group of children who stuttered.

In light of these findings, there is little empirical support for the hypothesis that language development and stuttering are linked to a common, underlying communication difficulty, at least in any significant number of young children. On closer examination of the research literature, methodological issues appear that may account for the view that children who stutter frequently have concomitant language difficulties. Several early studies of language ability in young children who stuttered did not consider socioeconomic status, potentially comparing young children who stuttered and were from lower or middle-income backgrounds with typically developing youngsters from university-based families (see Ratner, 1997). Furthermore, several past studies of language ability in young children who stuttered did not evaluate the children’s skills in light of normative expectations. Other studies reported higher than expected rates of concomitant language disabilities in young children who stuttered but evaluated children long after stuttering onset, or included children of very different ages. When relations between language ability and stuttering are examined long after stuttering onset, children may well have learned to adapt their expressive

Figure 1. Mean length of utterance, compared with normative expectations, for children whose stuttering persists or recovers. (From Watkins, R. V., Yairi, E., and Ambrose, N. G. [1999]. Early childhood stuttering. III. Initial status of expressive language abilities. Journal of Speech, Language, and Hearing Research, 42, 1125–1135. Reproduced with permission.)
language in various ways in order to limit or reduce stuttering events. Such studies may be interesting, but they ask very different questions from those addressed by investigations of language skills near the onset of stuttering. Any or all of these methodological choices could have considerable impact on findings pertaining to patterns and pathways of language acquisition in youngsters who stutter, and all could be in the direction that predict a less favorable performance for children who stutter in comparison with typically developing children.

In general, there is a growing consensus that language development is not particularly vulnerable in young children who stutter. Continued study of language strengths or challenges in conjunction with stuttering may reveal developmental asynchronies (e.g., perhaps early precocious language skill is a particular risk factor for stuttering, or perhaps accelerated language development in one domain, such as syntax or semantics, creates difficulties with fluency when proficiencies in other domains, such as motoric abilities, are less sophisticated). These possibilities await empirical study and will require detailed linguistic analyses to evaluate.

**Linguistic Influences on Stuttering**

There is ample evidence that stuttering events are influenced by linguistic variables. Brown (1945) was perhaps the first researcher to suggest linguistic influences on stuttering events with his groundbreaking report of apparent influences of a word’s grammatical form class (i.e., content versus function word) on stuttering loci in adults who stuttered. In brief, Brown reported that adults who stuttered were significantly more likely to be disfluent on content words (e.g., nouns and verbs) than on function words (e.g., prepositions and pronouns).

Since Brown’s seminal work, researchers have continually refined analyses in the study of linguistic influences on stuttering. We now know, for example, that young children are generally more likely to stutter on sentences of greater grammatical complexity than on sentences of less grammatical complexity (Logan and Conture, 1995; Yaruss, 1999). Furthermore, the content-function variable appears not to be the most relevant influence on stuttering loci; instead, stuttering events are significantly more likely on either a content word or a phrase-initial function word that precedes a content word than in other phrasal locations (Au-Yeung, Howell, and Pilgrim, 1998). The underlying influence here is thought to be the planning unit in language formulation, such that disfluencies are significantly more likely to occur at the beginning of a language planning unit, when remaining components of the unit continue to be refined for production.

These findings reveal that aspects of language planning, formulation, and production exert an influence on stuttering for children and adults. These findings support the view that linguistic variables are relevant in characterizing stuttering events. It is noteworthy, however, that linguistic factors appear to influence disfluencies in the same way for stutterers and nonstutterers alike. That is, linguistic variables such as grammatical complexity and the loci of stuttering tend to influence individuals whose disfluencies occur at typical rates in language production as well as individuals whose disfluencies are frequent enough to yield identification as a “stutterer.”

The domain of language is relevant in the study of early childhood stuttering. The majority of young children who stutter display expressive language abilities at or above normative expectations. In addition, a number of linguistic variables, such as the grammatical complexity of an utterance, exert an influence on the likelihood of a stuttering event. It may be informative for future investigation in this area of inquiry to move toward detailed and specific analyses of profiles of language strength and evaluation of synchrony versus asynchrony within and across developmental domains.

See also SPEECH DISFLUENCY AND STUTTERING IN CHILDREN.

—Ruth V. Watkins

**References**


Language of the Deaf: Acquisition of English

Effective English language skills are essential for the education of deaf children and for the integration of deaf individuals into the wider, hearing society. However, the average deaf student’s English language competence over the course of his or her schooling is limited. Although periodic reviews of nationwide achievement testing reveal some improvement in the average reading levels of deaf students over the past 30 years, the increase is small, and in English literacy skills most deaf students fall farther and farther behind their hearing peers over the course of their schooling. The average reading comprehension scores for deaf students in the United States rises from the second-grade level at age 9 to around the fourth-grade level at age 17 (de Villiers, 1992). Among white deaf high school graduates, only about 15% read above the sixth-grade level, and the percentage is only about 5% for those who are from African-American or Hispanic backgrounds (Allen, 1994).

The hearing child and the typical deaf child are in very different stages of language development when they reach the point of formal schooling and reading instruction. Normally hearing 5- or 6-year-old children have a speaking vocabulary of several thousand words and have mastered most of the complex syntax of English. Thus, for hearing children, the acquisition of reading is primarily learning to map printed English onto an existing knowledge of spoken English (Adams, 1990).

For the deaf child, the situation is quite different. Impairment of hearing has a negative effect on the acquisition of a spoken language from early in the child’s life, and all of the major milestones in normal language acquisition are considerably delayed. For example, around 6 months of age, normal-hearing infants begin producing the first approximations to consonant-vowel combinations, or syllables, over quite a wide range of speech-like sounds (so-called “marginal” babbling). A few months later “canonical” babbling emerges, in which the child produces a more restricted set of phonetic units in repetitive, rhythmic syllabic organization, and there is an abrupt decline in non-speech-like vocalizations. Canonical babbling increasingly uses the phonemes found in the adult input language, yet it makes no reference to any real-world objects or actions. However, the parents of these hearing infants treat their infants’ babbles as if they were meaningful, engaging in reciprocal “dialogues” and commenting on the infants’ utterances. For deaf infants, early vocalization and marginal babbling is not delayed, and they produce a similar range of speech-like sounds, which suggests that this stage is biologically driven. But canonical babbling is considerably delayed and appears to be deviant in both vocal quantity and quality (Mogford, 1993). So, hearing parents of deaf infants do not respond to their infant’s vocalizations as proto-communications, being more likely to ignore or talk through them. Thus, most deaf infants are already at a disadvantage toward the end of the first year of life, both in their ability to extract the phonemes of their spoken language from the input and reproduce them, and in the initial structuring of conversational dialogues in parent-infant interaction (Paul, 2001).

Phonology

Speech intelligibility is a persistent problem for deaf children with moderate to profound hearing loss (>60 dB loss in the better ear), particularly if the major hearing loss is in the higher frequencies of sound (>1500 Hz). Studies report from 50% to 80% of moderately to profoundly deaf children having either “very hard to understand” or “totally unintelligible” speech (Carny, 1986). Omission or distortion of consonant sounds is common and has a major impact on intelligibility (Osberger and McGarr, 1982). Many of the speech errors of deaf children reflect the phonological processes and constraints operating in normal speech development in young hearing children (e.g., consonant cluster reduction, fronting of place of articulation, voicing errors, or deletion of final consonants), but they persist in the speech of deaf children well into later childhood (Murphy and Dodd, 1995). Control of voice quality and intonation is also difficult for children with substantial hearing loss. Consequently, instances of consistently high or low pitch, nasalized speech, and rhythmic errors such as unusual breath groups and either misplaced syllabic stress or added syllables are common and produce a characteristic set of deaf “accents” in spoken English (Osberger and McGarr, 1982; Paterson, 1994; Murphy and Dodd, 1995).

Vocabulary

As measured by parental reports on the MacArthur Communicative Development Inventory (CDI), the average hearing child progresses rapidly from an expressive vocabulary of approximately 100 words at 18 months to 300 words at 2 years and 550 words at 3 years (Fenson et al., 1993). In contrast, the average deaf toddler with hearing parents produces about 30 words at 2 years and 200 words at 3 years, whether those words are spoken or
signed (Mayne, Yoshinago-Itano, Sedey, and Carey, 2000). Deaf children whose hearing loss was identified by 6 months of age and who have above-average cognitive skills fare best, showing a vocabulary spurt in the third year of life similar to that found in hearing children, though about 6 months later (Mayne et al., 2000). However, even these successful deaf children fall below the 25th percentile for hearing children on norms for the CDI at 30–36 months of age (see also Mayne, Yoshinago-Itano, and Sedey, 2000). Thus, the average 5- to 6-year-old deaf child is some 2 years behind hearing peers in vocabulary size when the child begins the task of learning to read.

Syntax

The English of deaf children also exhibits characteristic syntactic problems. For example, tense markers on the verb and many grammatical morphemes and function words (e.g., articles “a” and “the,” or copula and auxiliary verbs) are inconsistently provided or missing from most deaf children’s spoken or signed English when they reach the early grades of formal schooling. These aspects of English grammar continue to provide great difficulty for deaf children during the school years (Quigley and King, 1980; de Villiers, 1988). In the terms of generative grammar (Radford, 1990; Leonard, 1995), the functional categories Infllectional Phrase (IP) and Determiner Phrase (DP), which host the marking of tense for verbs and specification for nouns, may be incompletely specified in the grammar of deaf children. If so, one might expect problems also with a final functional category that structures the embedding of clauses, the Complementizer Phrase (CP) (de Villiers, de Villiers, and Hoban, 1994). Hearing children at the age of 5 or 6 have mastered a variety of multclause embedded sentence forms—especially temporal and causal adverbial clauses, complement structures, and relative clauses—that are essential for creating cohesion in narrative discourse and other extended conversation (de Villiers, 1988, 1991; Engen, 1994). However, in deaf children, clauses and sentences tend to be strung together with “and” or “then,” and complex embedded structures are usually missing or malformed in their English production and also are poorly comprehended (Engen and Engen, 1983; de Villiers, de Villiers, and Hoban, 1994; Engen, 1994; Berent, 1996). Thus the spoken, signed, or written English narratives of deaf students can often be characterized as a list of sentences each describing an event, but with little cohesion or coherence because the characters and events are not linguistically linked together by referential, causal, and temporal cohesion markers (de Villiers, 1991; Engen, 1994).

In summary, when deaf children reach the point of acquiring English literacy skills, they usually have a severely limited vocabulary and lack knowledge of the complex syntax of English that is critical for combining sentences together into cohesive text. Indeed, much of the deaf child’s English language learning comes from printed English, and the task of learning to read becomes one of both cracking the print code and learning the language at the same time. The better the child’s English language skills are before formal schooling, the easier the task of reading acquisition becomes.

New Developments

This pattern of delay and difficulty with language and literacy in deaf youngsters is mitigated by two technological advances that are transforming language intervention with deaf toddlers. The first is the implementation of universal newborn hearing screening programs, with the potential for identifying almost all infants with a significant degree of hearing loss within the first few months of life. By 2000, some two dozen states in the United States had mandated universal screening of hearing for infants prior to hospital discharge. Researchers in the state of Colorado demonstrated that identification of hearing loss prior to 6 months of age followed by effective ongoing intervention services and parent training programs was a major contributor to successful language outcomes for deaf children (Yoshinaga-Itano and Appuzzo, 1998a, 1998b; Yoshinago-Itano, Sedey, et al., 1998).

The second important advance is the development of multichannel cochlear implant technology. Electrodes implanted into the cochlear now bypass the hair cells and stimulate the auditory nerve directly. An external receiver and processor analyzes the incoming speech according to a predetermined strategy and transforms the complex pattern of sound frequencies and amplitudes into a corresponding pattern of electrical stimulation across a number of electrodes in the cochlear. Over the past 15 years, major developments in the complexity of the analysis that can be carried out in the externally worn speech processor (increasingly miniaturized by developing computer technology) first led to remarkable improvements in speech perception in postlingually deafened adults with implants. Now younger and younger profoundly deaf children who lost their hearing at or soon after birth and who would not benefit much from conventional hearing aids are receiving cochlear implants, some as young as 18 months of age (Niparko et al., 2000). Children who receive an implant early in life, followed by intensive auditory and speech training, can achieve speech intelligibility and conversational fluency that exceed the levels typically observed in profoundly deaf children who use hearing aids (Fryauf-Bertschy et al., 1997; Spencer, Tye-Murray, and Tomblin, 1998; Tomblin et al., 1999; Svirsky et al., 2000). However, there remains substantial individual variation in the degree of success of these implants with prelingually deaf toddlers (Pisoni et al., 2000). Although some children exhibit dramatic improvement in their speech perception and production, and may even acquire grammatical and vocabulary skills at a rate matching that of hearing children, others show only limited spoken language gains even after 3 or 4 years of implant use. The sources of this variability are still rather poorly understood but may include such factors as age at implantation, preimplant residual hearing, type of speech processor used, and success at tuning or “mapping” the
implant, as well as the educational and family environment (Fryauf-Bertschy et al., 1997; Niparko et al., 2000; Osberger and Fisher, 2000; Pisoni et al., 2000).

**Conclusion**

In summary, most children who are born with a moderate to profound hearing loss or who become deaf in the first few years of life suffer a pervasive disruption in their acquisition of all aspects of spoken English (and its signed forms). This leads to major difficulties in the children’s acquisition of literacy skills. Early identification of hearing loss and the immediate implementation of intervention strategies involving the best available technologies for amplification, and parental training in early language intervention in a rich interactional context, seem to offer the best outcome for these children in their acquisition of English.

—Peter A. de Villiers

**References**


**Further Readings**


**Language of the Deaf: Sign Language**

Hearing loss limits deaf children’s access to spoken languages, but deaf communities around the world acquire natural sign languages that create complete communication systems with the same subtlety and level of syntactic and semantic complexity as any spoken language. At the core of these communities are the 8%–10% of deaf children with deaf parents who are exposed to sign language from birth. But some deaf children with hearing parents are also exposed to a natural sign language through contact with native-signing deaf children and adults in educational settings. Thus, deaf children may be speech-delayed, but they are not necessarily language-delayed or disordered in any sense.

**Characteristics of American Sign Language and Other Natural Sign Languages**

This entry focuses on American Sign Language (ASL) because it is the natural sign language used in the United States and it has been the most extensively studied. However, many of the issues raised here about the unique properties of natural sign languages and the normal pattern of acquisition of those languages by deaf children exposed to complete and early input apply to all natural sign languages.

ASL and other natural sign languages are formally structured at different levels and follow the same universal constraints and organizational principles of all natural languages. Like the distinctive features of spoken phonology, a limited set of handshapes, movements, and places of articulation on the face and body distinguish different lexical signs. For example, in ASL the signs for SUMMER, UGLY, and DRY are produced with the same handshape and movement, but in different locations on the face (Bellugi et al., 1993). Just as in spoken languages, the syntactic rules of ASL operate on underlying abstract categories defined by their linguistic function, such as subjects and objects, or noun phrases and verb phrases. Furthermore, grammatical processes are recursive, embedding one phrase or clause within another (Liddell, 1980).

On the other hand, the visual-spatial modality of natural sign languages leads to several distinctive properties. Spoken languages are mostly sequential, in that the order of speech elements determines meaning. For example, temporal or adverbial modulations in meaning are expressed in spoken English by inflectional suffixes and prefixes that are added to the verb root. In contrast, multiple features of meaning are communicated simultaneously in sign languages: the place, direction, and manner in which signs are produced frequently add to or modulate the meaning of a sign. For example, ASL has evolved a system of simultaneous inflectional morphology on the verb that indicates person, number, distributorial aspect, and such temporal aspects as the repetition, habituality, and duration of the action. The single sign GIVE, for example, can be inflected to communicate the meanings “give to me,” “give regularly,” “give to them,” “give to a number of people at different times,” “give over time,” “give to each,” and “give to each over time” (Bellugi et al., 1993).

Second, sign languages use space as a grammatical and semantic device. For example, in ASL the noun referring to a particular person or object can be assigned (or indexed) to a location in space, typically to one or other side of the signer. Referring back to that place...
in space by pointing to it then acts as an anaphoric pronoun. Similarly, for verbs such as GO, GIVE, INFORM, or TEACH that involve directionality or movement in their meaning, the starting and end points of the sign and its direction of movement between points in space are used as an agreement marker on the verb to indicate the subject and recipient of the action (Wilbur, 1987; Bellugi et al., 1993).

Third, ASL makes extensive use of simultaneous nonmanual facial expressions and body movements as adverbial, grammatical, and semantic devices. Some of the facial expressions accompanying signed sentences seem to be expressions of intensity or attitude, such as pursing the lips or puffing out the cheeks, but others have obligatory syntactic roles. Raising the eyebrows and tilting the head forward slightly changes a declarative sentence into a yes/no question. Topicalized clauses such as relative clauses specifying information about a referent are marked by raised eyebrows and the head tilted back (Liddell, 1980). Finally, ASL has several ways to negate an utterance, and a nonmanual marker—a headshake with the eyebrows squeezed together—is used with or without the negative signs for NO and NOT to negate a clause (Humphries, Padden, and O'Rourke, 1980; Wilbur, 1987).

In reporting speech or action, the person whose perspective is to be taken is assigned to a location in space (Emmory and Reilly, 1995). Then the signer turns his body as if signing from the perspective of that location and signs what the person did or said. This “role shift” is analogous to direct speech in spoken languages. It is accompanied by a break in eye gaze away from the conversational partner and the use of “first-person” pronouns in the role of the character involved. Reporting action also uses a role shift of the signer’s body in space, but it employs different pronoun usage and facial markers to indicate that the actor’s perspective is being taken (Emmory and Reilly, 1995).

Finally, like some spoken languages, ASL incorporates classifiers, linguistic markers that identify such features as the size, shape, animacy, and function of objects. In ASL different handshapes encode these properties of the objects and are incorporated into movement verbs (Wilbur, 1987; Schick, 1990). The most iconic of the classifier handshapes are those that reflect the size and shape of the object referred to. So, a single handshape depicts all medium-sized cylindrical objects, such as a cup, a small tube, and a vase. Other classifiers are more abstract, representing classes of objects that do not resemble each other (or the classifier sign) in size or shape. Thus, one classifier is used to represent all vehicles, including boats and bicycles. Some classifier signs can serve as pronouns in sentences (Humphries, Padden, and O'Rourke, 1980).

**Acquisition of ASL by Native-Signing Children**

Deaf children exposed at an early age to a sufficiently rich input acquire a full natural sign language as effortlessly and rapidly as hearing children acquire their native spoken language. Deaf babies “babble” in sign at about the same age as their hearing peers babble in speech, repeating the handshape or movement components of signs in a rhythmic fashion (Pettito and Marentette, 1991). Just as in canonical babbling the child’s phonetic repertoire comes to be restricted to that of the child’s native language, so too the sign-babbling deaf child shifts to incorporating only the restricted set of handshapes and movements found in the input sign language. Notably, the phonetic units found in “canonical” man-

![Figure 1](From Bellugi, U., et al. [1993]. The acquisition of syntax and space in young deaf signers. In D. Bishop and K. Mogford [Eds.], Language development in exceptional circumstances. Hillsdale, NJ: Erlbaum. Reproduced with permission.)
ual babbling are the same ones later used in the first meaningful signs (Pettito, 2000).

Although there is considerable variation among children, the first signs with a clear reference may emerge a month or two earlier in deaf children than the first spoken words of hearing children. Motor control over the hands and arms sufficient to produce recognizable signs develops a little ahead of control over the vocal articulators (Bonvillian, 1999). Early vocabularies in sign or speech refer to the same categories of objects and actions: the significant people, animals, objects, and actions in the common environment of most toddlers, hearing or deaf. Some signs in ASL are iconic in that they resemble the referent. For example, the sign for CAT is made by stroking the side of the upper lip to indicate whiskers. However, while iconicity may facilitate the comprehension of some signs, semantic domain and phonological complexity seem to be stronger determinants of early sign productions (Bonvillian, 1999).

The phonological properties of early signs are affected by motor control and perceptual salience, as well as by linguistic constraints (Conlin et al., 2000; Marentette and Mayberry, 2000). For example, sign location seems to be the most accurately produced feature of early signs because the place of articulation is perceptually more salient than the handshape and manner of movement and requires less fine motor control. Signing children produce a pattern of regular substitutions of phonetic elements in their early signs, just like the phonological substitutions in hearing toddlers’ early spoken words (Bellugi et al., 1993).

By around 2 years of age, children begin to master the pronoun system, in ASL a system of pointing gestures to the self (“me”), to one’s conversational partner (“you”), or to indexed referential spaces (“he,” “she,” “it”). Despite the iconic transparency of first- and second-person pronouns, signing children follow the same developmental timeline and exhibit the same problems as speaking children do in learning pronouns that shift reference, depending on who is the speaker and addressee. They even tend to make the same reversal error in early acquisition of “I/me” and “you,” using a point facing outward toward their conversational partner when referring to themselves, and pointing to themselves when referring to “you” (Pettito, 1987). More abstract pronouns that involve pointing to an indexed referent assigned to a location in signing space are the last acquired, around age 5 (Hoffmeister and Wilbur, 1980).

In two- and three-word signed utterances, semantic relationships between sentence elements emerge in a reliable order of acquisition that seems to be determined by the conceptual development of the child: reference to the existence or disappearance of objects, then action relationships, and finally state relationships (properties, possession, or location of objects), just as in early spoken language acquisition (Brown, 1973; Newport and Ashbrook, 1977).

Although much of the syntax of ASL is conveyed through spatial morphology rather than word order, young signing children begin by producing uninflected root forms of these verbs and use more fixed word orders in their early sentences to express grammatical and semantic relationships, typically the most frequent sign order seen in the adult input (Newport and Meier, 1985). The use of agreement morphology on these verbs emerges over the period between age 2 and 5 years, being mastered first for referents that are present in the discourse context and only later for absent referents indexed in space (Newport and Meier, 1985). The pattern of development and the errors are predicted not by iconicity, the fact that the path of the action is clearly traced in space, but by a linguistic model of morphological complexity observed in spoken languages (Newport and Meier, 1985; Slobin, 1985). Just as speaking children overgeneralize inflectional rules like the regular -ed past tense (e.g., “helded” for “held”) or use noncausative verbs in a causative sense (e.g., “He failed me down”), signing deaf children overgeneralize spatial agreement marking to verb signs in ASL that cannot be used in that way. So, verbs like SAY or LIKE, which cannot be directionally marked, are sometimes extended toward the object, and verbs like DRINK and EAT may be signed as if coming from an indexed referent in space, not in their required location on the signer (Bellugi et al., 1993).

Classifiers first emerge at around age 3 but are mastered over a long period of time, with some forms still giving children trouble at age 6 or 7. Size and shape classifier handshapes resemble their referents more closely and are acquired first (Schick, 1990). Mastery of the more abstract classifiers depends not only on mastery of the syntax and semantics of ASL, but also on the child’s conceptual development in classifying such objects as cars, boats, and bicycles all together as a functional class.

Despite our natural attention to affective markers in facial expressions, nonmanual syntactic and semantic markers can be difficult for both first and second language learners of ASL to acquire (Reilly, 2000). The marker itself may appear early, but its full linguistic use can take several years to master. Thus the negative headshake is first used by native-signing deaf children in the second year to negate signs, but young children experience difficulty in timing this nonmanual marker to coincide with the correct manual signs, and they may ungrammatically extend the headshake across more than one clause or sentence (Anderson and Reilly, 1997).

Thus the same developmental processes and constraints apply to the natural acquisition of ASL and other sign languages as to the normal development of any spoken language. The pattern of acquisition is primarily dictated by the linguistic and cognitive complexity of forms, not by their iconicity in the visual mode of the language.

**Natural Sign Language Versus Artificial Signed Versions of Spoken Language**

In several countries, signed versions of the native spoken language have been created by educators. Unlike natural
sign languages, they are typically signed simultaneously with the spoken language as a form of sign-supported speech. In the United States there are several versions of manually coded English (MCE). The most widely used of these are Signed English, Signed Exact English, and Seeing Essential English. These all use lexical signs from ASL and English word order, but they vary in the degree to which created signs encode all of the function words or derivational and inflectional morphology of English.

The educational value of these MCE systems over ASL and oral or written English is hotly disputed. They have two primary drawbacks. First, because signs take longer to produce than words, signing a sentence in a linear sequence following that of English takes much longer than speaking the same sentence in English. The simultaneous speaker-signer either has to slow down her speech to an unnatural extent or has to leave out aspects of the signed portion of the message. Thus, deaf children receive either an incomplete and ungrammatical language input or a simplified one, so that their exposure to complex English syntax is artificially reduced.

Second, several universal features of natural sign languages that have evolved to allow effective and rapid communication of meaning in a visual-spatial mode are not incorporated into MCE systems. These features include the use of space as a grammatical and semantic device, simultaneous morphology, and nonmanual linguistic markers. Indeed, in several respects MCE systems directly violate these universal principles and so can be very confusing for deaf children who have been exposed to ASL (Johnson, Liddell, and Erting, 1989). Deaf children may then naturalize MCE so that it more closely conforms to ASL’s use of space and simultaneity (Suppala, 1991).

**Going Beyond the Input**

Many deaf children are exposed to incomplete versions of ASL, from deaf parents who are not native signers or hearing parents still learning ASL. In this input morphological rules in particular may be inconsistently used. However, children systematize their sign language to make consistent rules out of what are only statistical regularities in their parents’ signing (Newport, 1999; Newport and Aslin, 2000). For example, spatial agreement markers on verbs of motion that were correct only about two-thirds of the time in the parental input were used correctly more than 90% of the time by their children at age 7. The children create a more regular rule out of inconsistently used morphology, thus going beyond the input to acquire a more complete form of ASL.

Even with no conventional signed input, deaf children of hearing parents in oral environments create rich gestural systems, although these are not complete languages (Goldin-Meadow, 2001). Over a longer period of time and several generations of deaf signers, a more complete sign language may emerge. This phenomenon is seen in the creation of a new Nicaraguan Sign Language out of many home gestural systems by deaf children brought together from isolated villages into a centralized school for the deaf in Nicaragua in the early 1980s. Motivated by the pressure to communicate among themselves, the deaf students evolved a more and more complex sign language, with each generation of students inheriting a more complex form of the language and then elaborating it. Several researchers have identified the emergence of many of the apparently universal features of natural sign languages in the evolution of Nicaraguan Sign Language over the past 20 years, among them componentiality of lexical signs, simultaneity of meaning expression, nonmanual syntactic markers, and the use of space for grammatical purposes (Kegl, Senghas, and Coppola, 1999; Senghas, 2000).

**A Critical Period for Sign Language Acquisition**

There seems to be a sensitive or even critical period in early childhood in which exposure to a relatively complete sign language must take place. Deaf individuals exposed to ASL for the first time in late childhood or adolescence are less proficient ASL users in adulthood than those who learn ASL in the first few years of life, even though they may have been using ASL as their primary means of communication for decades (Fischer, 1998; Mayberry and Eichen, 1991; Newport, 1991).

**ASL and Finger Spelling**

There are some influences on ASL from the surrounding dominant English language culture, for example the importation of loan words from English that are finger-spelled using the manual alphabet to represent English letters. Finger-spelled vocabulary is highly selective, almost always nouns and rarely verbs. In everyday conversations these finger-spelled words are frequently the names of people and places, but in educational settings finger spelling is used to represent technical words and concepts for which there is not a natural sign in ASL. Padden (1998) has argued that this is similar to the way in which most if not all spoken languages import foreign vocabulary words. Indeed, deaf children of deaf parents who cannot yet read or write English begin to produce finger-spelled words as “signs” before they learn to connect them to English orthography.

Padden and Ramsey (1998) suggest that finger spelling may interact with ASL skills in providing deaf children with better access to English literacy learning. They find that deaf teachers using ASL in the classroom actually finger-spell more words than hearing teachers. Most important, these deaf teachers effectively chain together the different expressive modes of communication, switching from finger spelling to printed text to finger spelling, or from ASL sign to finger spelling to text, in order to facilitate the connection between text and meaning and to develop better print decoding skills. In Padden and Ramsey’s study, the deaf students with better ASL and finger-spelling skills developed better English-reading skills.
**ASL and English Literacy**

The extent to which the development of phonological decoding from print to spoken phonemes is necessary for fluent reading in deaf students is still in dispute. However, skill in ASL does not interfere with learning to read printed English. Rather, it is a strong independent contributor to reading comprehension levels for deaf children in educational settings using sign language, even when controlling for having deaf or hearing parents (Hoffmeister et al., 1997; Strong and Prinz, 1997). Having a fluent sign language can facilitate learning to read in several ways: by increasing the children’s comprehension of the instructional process (if the teacher is also fluent in ASL), increasing the number of semantic concepts the child understands, developing extended discourse skills that are critical to early reading, and fostering metacognitive skills such as communication monitoring and planning (Nelson, 1998).

**Conclusion**

The visual-spatial nature of sign languages—the fact that they are articulated with the hands and perceived through the eyes—does not relegate them to the realm of pantomime and gesture. Natural sign languages are as subtle and complex as any spoken language and are structured according to universal linguistic principles. Deaf children exposed to a complete and consistent natural sign language early in childhood acquire the language normally, following the same stages and learning processes as are observed in hearing children acquiring their native spoken language.

—Peter A. de Villiers and Jennie Pyers

**References**


Linguistic Aspects of Child Language Impairment—Prosody

In linguistics, “prosody” refers to sound patterns in language involving more than a single segment or phoneme. Since the early 1980s, the study of prosody has blossomed, both in linguistics and in the allied areas of computer speech analysis and synthesis, adult sentence processing, infant speech perception, and language production. The study of prosody has provided insight into the word and sentence productions of young children with normally developing language and language disorders. In particular, prosody has proven a useful tool for examining children’s “deviant” utterances; that is, utterances that deviate from what we would expect from an adult speaker with normal speech and language. For example, a child’s production of “blue” as “bu” will be considered a deviant utterance for purposes of this discussion.

The article begins with a brief overview of two aspects of prosody, syllable shape and meter, that have been the focus of many studies of child language production. It then provides examples of some recent studies that demonstrate effects of syllable shape and meter on deviant productions of children with normal and disordered language. To foreshadow the general finding across the studies presented here, we state that syllable shapes and metrical patterns that are frequent across the world’s languages or in the language the child is learning are most resistant to deviations.

Beginning with syllable shape, it has been noted that all languages of the world have syllables comprising a consonant plus a vowel (CV). Only some languages allow additional syllable shapes, such as V, VC, CVC, CCVC, and so on. In linguistic terms, the CV syllable shape is said to be “unmarked.” Even when languages allow syllable shapes other than CV, the shapes are often restricted. For example, Japanese allows only CVC syllables ending in /n/.

Furthermore, in languages like English that allow consonant clusters (syllables of the shape CCVC, CVCC,
etc.), these clusters generally conform to a sonority sequencing principle (e.g., Hooper, 1976). Briefly, sonority may be conceptualized as the openness of the vocal tract for a particular segment. Vowels are the most sonororous; glides are less sonororous, followed by liquids, nasals, fricatives, and stops. Clements (1990) argues that an ideal syllable structure is one in which a sequence of segments increases from the onset to the vowel with no or a minimal decline from the vowel to the coda. English word-initial consonant clusters such as /pr/ and /kw/, which comprise a stop plus liquid or glide, are consistent with this principle, as are English word-final clusters such as /nt/ and /nt/.

The foregoing discussion of syllable shapes concerns what is allowed in words of a language. It is also important to note that even languages that allow a variety of syllable shapes nevertheless have strong statistical tendencies toward particular shapes. For example, English CVC words are very likely to end in /t/ and very unlikely to end in /d/. These statistical tendencies are the subject of a growing interest in researchers studying prosody and its role in child language production.

Let us now turn to meter. In many languages, multi-syllabic words exhibit a characteristic stress pattern. For example, the majority of words in English have the pattern found in “apple” and “yellow,” that is, a stressed or strong syllable followed by an unstressed or weak syllable. Stressed syllables are louder, longer, and higher in pitch than unstressed syllables. The frequency of the strong-weak word pattern is consistent with the observation that the basic unit of stress in English is a trochaic (strong-weak) foot, with a foot defined as a grouping of a single strong syllable plus adjacent weak syllables.

Feet not only explain the dominant stress pattern of words in a language, they also help us to understand how lexical words like nouns and verbs combine with grammatical words like determiners and auxiliary verbs in phrases. For example, when we say phrases like “drink of water” and “pick a card,” we tend to combine the grammatical words (in this case “of” and “a”) with the preceding strong syllable, even though these words belong syntactically with the following word (in a prepositional phrase and noun phrase, respectively). That is, English speakers tend to create trochaic feet whenever they can, giving the language a characteristic metrical pattern.

The discussion up to this point has revealed that languages of the world and particular languages are biased toward specific syllable shapes (e.g., CV) and meters (e.g., trochaic). Beginning with syllable shape, let us now consider how these prosodic biases affect the productions of children with normally developing and disordered language. Two of the most frequent syllable shape deviations from a standard target produced by children are final consonant or coda deletion and consonant cluster reduction. With respect to coda deletion, this phenomenon has been viewed as one in which the speaker is resorting to the most common syllable shape, CV. However, recent studies indicate that there are significant differences in the rate at which children omit different codas in different prosodic environments. Zamuner and Gerken (1998) reported that normally developing 2-year-olds produced more codas and more coda types on nonsense words when the coda occurred in a stressed syllable (either on a monosyllabic item or an item with a weak-strong stress pattern, e.g., /mab/b/). Zamuner (2001) discovered that children from the same population produced obstruent codas, which are more frequent in English, sooner than sonorant codas on CVC nonsense words. She also found that the same coda was produced less frequently when it occurred in nonsense names exhibiting less frequent biphone (e.g., CV, VC) than more frequent bifonhes.

With respect to consonant cluster reduction, several studies have shown a role for syllable shape, and in particular sonority sequencing, in this phenomenon (e.g., Barlow and Dinnsen, 1998; Ohala, 1999). These studies have revealed that children with normal language and language disorders were more likely to produce the least sonorous consonant of an initial cluster and the least sonorous consonant of a final cluster. That is, they produced CV sequences that were closer to the ideal syllable shape suggested by Clements (1990).

Turning now to the role of meter in language production, several researchers have noted that children are more likely to omit weak syllables from the beginning of words like “giraffe” and “banana” and, more generally weak syllables that do not belong to a trochaic foot (e.g., Wijnen, Krikhaar, and Den Os, 1994). The bias to produce trochaic feet has also been observed at the level of sentence production, where the determiner “the” is more likely to be preserved in a sentence like “He pats the zebra” (“pats the” forms a trochaic foot) than “He brushes the bear” (the syllabic verb inflection makes the formation of a trochaic foot containing “the” impossible; Gerken, 1996). Thus, the effects of prosody are not restricted to what have traditionally been considered phonological deviations but extend to morphosyntactic deviations as well. It is interesting to note that not all languages show as strong a bias toward trochaic feet as English does. For example, Spanish has many words like “banana,” which exhibit a weak-strong-weak pattern. Spanish-learning children have been shown to produce determiners at an earlier age than their English-learning counterparts, again suggesting a role for prosody in children’s morphosyntactic development (Lleó and Demuth, 1999).

At least some of children’s weak syllable omissions appear to occur during late stages of language production rather than during utterance planning, as evidenced by work by Carter (1999). Normally developing 2-year-olds and older children with language impairment produced sentences like “He kissed Cassandra” and “He kissed Sandy.” Note that the former sentence type was frequently produced with the first syllable of the name omitted (Cassandra → Sandra). Acoustic measurements revealed that, even though the two types of sentences contained the same number of overtly produced syllables (four), children produced the first sentence type with a longer duration, suggesting that they reserved a
timing slot for the syllable they eventually omitted. One possible source of weak syllable omissions is a lack of complete control over the motor sequences involved in producing trochaic vs. weak-strong feet (Goffman, 1999).

Finally, several studies have revealed joint effects of syllable shape and meter on deviant utterances. In the Zamuner and Gerken study discussed above, children showed different rates of coda preservation for strong and weak syllables. Ohala (1998) found that young children with normal language were less likely to reduce word-medial consonant clusters in words with a strong-strong stress pattern. In a study of weak syllable omission in young children with normal language, Kehoe and Stoel-Gammon (1997) noted more omissions of the middle syllable of words like “elephant,” which exhibit a strong-weak pattern, if the syllable began with a sonorant consonant. Carter (1999) found that adults with a variety of types of aphasia were more likely to omit word-initial weak syllables with V and VC syllable shapes than CV shapes. It seems likely that such results would be found in children with normal and disordered language as well.

In summary, linguistic studies of canonical syllable shapes and metrical patterns across languages and within particular languages provide the tools for fine-grained analyses of deviant forms produced by children with normal and disordered language. The results of these analyses feeds a growing consensus that those forms that are very frequent in languages of the world or in the child’s target language are generally more robust and less susceptible to deviations from the accepted standard. Further research is needed to reveal the mechanism underlying prosody’s clear effect on language production (see also PROSODIC DEFICIENCES).

—LouAnn Gerken

References


Further Readings


Melodic Intonation Therapy

Among the many published approaches for the treatment of aphasia, melodic intonation therapy (MIT) is one of the few techniques whose clinical effectiveness has been established by peer review (American Academy of Neurology, 1994). The effectiveness of the program is based on the specific guidelines for patient candidacy, its formalized protocol, and variety of reports testifying to improved communication competence following MIT. After evaluating the available evidence, the American Academy of Neurology considers the program to be promising when administered by a qualified speech-language pathologist.

The guiding principles and procedures associated with MIT were set forth in the early works of Albert, Sparks, and Helm (1973), Sparks, Helm, and Albert (1974), and Sparks and Holland (1976). More recent descriptions of the program can be found in Helm-Estabrooks and Albert (1991) and Sparks (2001). Generally, three principles form the conceptual foundation for MIT. First, in most of the population, the right cerebral hemisphere mediates music and speech prosody. Second, the right hemisphere is preserved in most individuals with aphasia, and as a result, singing abilities are generally preserved even in the most severe cases of aphasia. Third, the preserved musical and prosodic capabilities of the right hemisphere can be exploited to rehabilitate language production in patients with aphasia.

The goals of MIT are to facilitate some recovery of language production in severely nonfluent speakers with poorly articulated or severely restricted verbal output. Good candidates have poor repetition but at least moderately preserved to essentially normal language comprehension. Attempts at self-correction are evident. They are emotionally stable, if sometimes depressed, and highly motivated to improve their speech. A coexisting buccofacial apraxia is usually observed, as well as right hemiplegia that is greater in the arm than leg. The program therefore seems to be particularly suited for patients with Broca’s or mixed nonfluent aphasia with accompanying apraxia of speech (Tonkovich and Peach, 1989; Square, Martin, and Bose, 2001). These characteristics also generally exclude patients with Wernicke’s, transcortical, or global aphasia.

The initial computed tomographic profile for good candidates included a large lesion in Broca’s area extending superiorly to the left premotor and sensorimotor cortex for the face and deep to the periventricular white matter, putamen, and internal capsule. The lesion also typically spared Wernicke’s area and the temporal isthmus. No lesions of the right hemisphere were detected; this evidence was used to support the preservation of melodic functions in these patients (Naeser and Helm-Estabrooks, 1985). Naeser (1994) subsequently identified two important areas in the subcortical white matter that appeared to have an important role regarding recovery of spontaneous speech. Lesions of good responders involved no more than half of the total area, including the medial subcallosal fasciculus and the middle one-third of the periventricular white matter. The extent of lesion in cortical language areas, including Broca’s area, could not be used to discriminate among individuals who responded well or poorly to MIT. Lesions may have involved Wernicke’s area or the subcortical temporal isthmus, but when they did, they involved less than half of those areas.

During the beginning stages of an MIT program, emphasis is placed on the production of syntactically and phonologically simplified phrases and sentences that gradually increase in complexity throughout the course of the program. Ideally, language materials are thematically related and relevant to the patient’s daily needs and background. A large corpus of materials is recommended to vary the stimuli from session to session and to decrease practice effects. It is debatable whether the use of supplementary pictures or written sentences is appropriate (Helm-Estabrooks and Albert, 1991; Sparks, 2001). Frequent treatment, perhaps twice daily, is essential, but when unattainable, family members might be used to assist with the program (Sparks, 2001).

MIT focuses on three elements represented in the spoken prosody of verbal utterances: the melodic line or variation in pitch in the spoken phrase or sentence, the tempo and rhythm of the utterance, and the points of stress for emphasis. The intoned pattern has a range of only three or four whole notes that is selected from several reasonable speech prosody patterns for the target sentence. Tempo is slowed by syllable lengthening; phrase accuracy appears to be best when syllable durations approximate 2.0 s per syllable. The effects of this tempo are most pronounced when patients are required to intone utterances independent of the clinician (Laughlin, Naeser, and Gordon, 1979). Rhythm and stress are exaggerated by elevating intoned notes and increasing loudness. Clinicians tap out and further reinforce the rhythm and stress of the utterances using...
the patient’s hand. The emphasis on slow tempo, precise rhythm, and distinct stress appears to facilitate the processing of the structure and the articulation of the intoned utterances.

The MIT program consists of four levels. In level I, the clinician hums a melody pattern within the three- to four-note range and aids the patient in tapping the rhythm and stress of the stimulus melody to establish the process of intoning melody patterns with hand tapping. Level II requires the patient to tap and repeat the clinician’s production of the intoned utterance and to respond to a probe question eliciting an intoned repetition of the intoned utterance. Hand tapping is not used in response to probe questions. The clinician provides assistance by intoning the utterance in unison with the patient and then fading his participation so that the patient subsequently intones the utterance on his own. In level III, unison intoning of the utterance is followed by immediate fading of the clinician’s participation. The patient then produces the target utterance following an enforced delay after the clinician presents it. Finally, the patient gives an appropriate intoned response to an intoned probe question from the clinician. A backup procedure is introduced at this level to provide the patient an opportunity to correct errors. The backups consist of repeating the previous step and attempting the failed step again, and as such constitute an “indirect” approach to correcting errors. The goal of level IV is normal speech prosody. Latencies for delayed repetition are increased, and the training sentences become more complex. A technique called Sprechgesang (speech-song) is used in the transition to speech prosody. In this technique, the constant pitch of the intoned words is replaced by the variable pitch of speech while retaining the tempo, rhythm, and stress of the intoned sentence. Unison production of the target sentence in Sprechgesang is followed by fading, delayed spoken repetition using normal speech prosody, and production using normal prosody in response to a probe question with normal speech prosody.

MIT uses a scoring method where values of 2, 1, or 0 can be obtained. Full scores (i.e., 1 for items with no backups, 2 for items with backups) are assigned to successful responses, while partial scores (i.e., 1) are assigned to responses that require a backup where available. No score is assigned to unsuccessful responses following multiple attempts. The average score for three sessions must be higher than the average score of the three previous sessions for the participant to remain in the program. An overall score of 90% or better for five consecutive sessions is required to advance from one level of MIT to the next.

The neurophysiological model offered by the developers of MIT to account for its effectiveness has been controversial since it was first proposed. Berlin (1976) stated that the evidence linking the right hemisphere to the interpretation of nonverbal acoustic processes like music is insufficient to conclude that MIT activates the right hemisphere in some way to control motor speech gestures. Instead, he suggested that good candidates for MIT might have an intact left primary motor area that is deprived of input from the damaged left Broca’s area. Improved speech production might then result from transcallosal input to left hemisphere speech motor centers arising from the MIT-activated right hemisphere homologue of Broca’s area. An alternative explanation involved input from a disconnected intact left Broca’s area to an intact left primary motor area via a transcallosal pathway involving the right hemisphere homologues to these areas.

Belin et al. (1996) used positron emission tomography to investigate recovery from nonfluent aphasia following treatment with MIT. Changes in cerebral blood flow were measured while the participant listened to and repeated simple words, and during repetition of intoned words. Abnormal activation of right hemisphere structures homotopic to those normally activated in the intact left hemisphere was observed during the simple word tasks performed without intoning, while word repetition with intoning reactivated essential motor language zones, including Broca’s area and the adjacent left prefrontal cortex. Belin et al. concluded that MIT is more strongly associated with exaggerated speech prosody than with singing and therefore recruits language-related brain areas of the left hemisphere rather than right hemisphere areas.

Boucher et al. (2001) investigated whether the processing of melodic contours in music applies similarly to the processing of speech prosody. According to these authors, melody is associated with musical tone and rhythm. Tonal elements include pitch, timbre, and chord and correspond to intonation in speech. Musical rhythm refers to the timing distribution of tonal elements and is comparable to the stress points of speech. Although there is support for right hemisphere processing of intonation, Boucher et al. (2001) provide evidence that the left hemisphere is involved in the processing of rhythm, and consequently question whether melody-based interventions such as MIT facilitate speech production because of right hemisphere contributions. Following interventions in two speakers with nonfluent aphasia using stimuli emphasizing tone or rhythm in varying conditions, equal or greater success in responding was found for conditions emphasizing rhythm than for conditions emphasizing melodic intoning. Boucher et al. concluded that the right hemisphere explanation for the facilitating effects of MIT could not be supported strongly.

—Richard K. Peach

References


Memory and Processing Capacity

Research on the role of working memory in language disorders has stemmed mainly from the phonological loop model (e.g., Baddeley, 1986; Gathercole and Baddeley, 1993) or the capacity theory of comprehension (e.g., Just and Carpenter, 1992). These models differ in their conception of working memory and in the paradigms typically used to assess this construct (cf. Montgomery, 2000a); however, a central premise of both frameworks is that there is a limited pool of operational resources available to perform computations, such that processing and storage of linguistic information is degraded when demands exceed available resources. Numerous investigations based on these two approaches have demonstrated an association between working memory capacity and normal language functioning in children and adults. In young children, individual differences in phonological working memory predict vocabulary development and are related to differences in word repertoire, utterance length, and grammatical construction use (e.g., Gathercole and Baddeley, 1990b; Adams and Gathercole, 2000). School-age children’s performance on working memory measures is significantly correlated with spoken language comprehension as well as with reading recognition and comprehension (e.g., Gaulin and Campbell, 1994; Swanson, 1996). Working memory capacity predicts a number of verbal abilities in adults, including reading comprehension levels, understanding of ambiguous passages and syntactically complex sentences, and the ability to make inferences (e.g., King and Just, 1991; Carpenter, Miyake, and Just, 1994).

Investigators have examined short-term or working memory abilities in children with varying profiles of language and cognitive deficits, including children with Down syndrome, Williams’ syndrome, Landau-Kleffner syndrome, learning disabilities, and specific language impairment (SLI). Of special interest are children with SLI, who demonstrate significant language deficits in the absence of any clearly identifiable cause such as mental retardation or hearing loss. One theoretical camp views SLI in terms of limited processing capacity. There are various formulations of limited capacity accounts of SLI, including hypotheses about specific deficits in phonological working memory and hypotheses regarding more generalized difficulties in information processing and storage that affect performance across modalities (cf. Leonard, 1998). Difficulties discussed here are limited to poor nonword repetition, reduced listening span, and poor serial recall.

Children with SLI exhibit deficits in nonword repetition, a paradigm that has been used extensively by Baddeley and colleagues (and others as well) as a measure of...
phonological working memory (Gathercole and Baddeley, 1990a; Montgomery, 1995; Dollaghan and Campbell, 1998; Edwards and Lahey, 1998; Ellis Weismer et al., 2000; Briscoe, Bishop, and Norbury, 2001). Nonword repetition has proved to be useful clinically as a culturally nonbiased measure for distinguishing between children with and without language disorders. In one of the initial investigations of nonword repetition in SLI, Gathercole and Baddeley (1990a) concluded that children with SLI demonstrate significantly poorer phonological working memory than controls matched on nonverbal cognition or language level (however, see van der Lely and Howard, 1993; Howard and van der Lely, 1995). The findings of Gathercole and Baddeley (1990a) were replicated by Montgomery (1995), who similarly interpreted his results as indicating that children with SLI have reduced phonological memory capacity.

Other studies have sought to determine whether difficulties with nonword repetition reflect cognitive processes other than working memory deficits (Edwards and Lahey, 1998; Briscoe, Bishop, and Norbury, 2001). After a thorough investigation of possible explanations for nonword repetition deficits in SLI, Edwards and Lahey (1998) concluded that neither auditory discrimination nor response processes could account for the difficulties. Instead, they attributed the deficits to problems in the formation or storage of phonological representations in working memory. Children with SLI usually do not differ from normal language peers in their ability to repeat short, simple nonwords; rather, breakdowns on nonword repetition tasks typically occur on the most complex stimuli (Ellis Weismer et al., 2000; Briscoe, Bishop, and Norbury, 2001). When children with SLI were compared with children with mild to moderate hearing loss, both groups showed similar difficulty with longer nonwords, but children with SLI also displayed deficits on digit recall and were more negatively affected by phonological complexity (Briscoe, Bishop, and Norbury, 2001). These investigators concluded that auditory perceptual deficits are not sufficient to explain the range of language and literacy difficulties observed in children with SLI and suggested that some kind of processing capacity limitation underlay their language deficits.

Several genetic investigations of developmental language disorder have examined phonological memory as indexed by nonword repetition. Bishop, North, and Donlan (1996) administered a nonword repetition test to participants in a study of twins with language impairment. Children with persistent language impairment as well as those with resolved language impairment exhibited significant deficits in nonword repetition. Comparison of nonword repetition performance in monozygotic and dizygotic twin pairs revealed a significant heritability component. Based on these results, Bishop et al. suggested that deficits in nonword repetition provide a phenotypic marker of heritable forms of developmental language disorder. Bishop and colleagues (1999) replicated the earlier results and found that nonword repetition gave high estimates of group heritability. This measure was a better predictor of low language scores than was a measure of auditory processing (Tallal’s Auditory Repetition Test). Tomlin and colleagues (2002) recently investigated candidate genes associated with developmental language disorder, testing for associations between candidate loci in a sample of 476 children and their parents. A two-stage approach was used to search for loci associated with language disorder, as diagnosed by standardized tests of listening and speaking or by a measure of phonological memory (nonword repetition). Preliminary results were suggestive of an association of CFTR (a marker on chromosome 7) with both the phonological memory and spoken language phenotypes.

Another paradigm widely employed in research on the association between language and working memory abilities uses a listening/reading span task (e.g., Daneman and Carpenter, 1980). The person is required to perform two tasks concurrently (involving processing and storage), such as making true/false judgments about sentences and recalling the last word in each sentence following the presentation of all sentences in a set. The number of sentences within a set increases throughout the task in order to assess memory span. Ellis Weismer, Evans, and Hesketh (1999) found that children with SLI evidenced limitations in verbal working memory compared to age-matched controls, based on their performance on a listening span task developed by Gaulin and Campbell (1994). Findings primarily pointed to quantitative differences between the groups involving reduced capacity for the children with SLI; however, there were some indications of qualitative differences in terms of distinct patterns of word-recall errors and different patterns of associations between working memory and performance on language and nonverbal cognitive measures. Montgomery (2000a, 2000b) examined the relation between working memory and sentence processing in children with SLI. Using an adaptation of Daneman and Carpenter’s listening span task, he demonstrated that children with SLI exhibit reduced capacity under dual-load conditions. Performance on the listening span measure was significantly correlated with performance on an off-line sentence comprehension task but not with on-line sentence processing. Montgomery concluded that the slower real-time sentence processing in children with SLI was primarily a function of inefficient lexical retrieval operations rather than limitations in working memory; however, he posited that their difficulties with off-line sentence comprehension tasks were related to difficulties coordinating the requisite processing and storage functions, revealing limitations in functional working memory capacity.

Serial memory deficits in children with SLI have been documented by Gillam and colleagues (Gillam, Cowan, and Day, 1995; Gillam, Cowan, and Marler, 1998). The initial study employed a suffix effect procedure in which a spoken list to be recalled was followed by a “suffix” (nonword item) that was not to be recalled. The suffix had a disproportionately negative effect on recency recall for the children with SLI when strict serial position criteria for scoring were imposed. The subsequent investi-
gation by Gillam et al. sought to determine the nature of working memory deficiencies in children with SLI, using a modality effect paradigm in which input modality, rate of input, and response modality were manipulated. To control for differences in capacity across the groups, trials were administered at a level consistent with each child’s working memory span. Children with SLI and controls demonstrated traditional primacy, recency, and modality effects and similar performance when audiovisual stimuli were paired with spoken responses. However, children with SLI exhibited reduced recency effects and poor recall when visually presented items were paired with pointing responses. The investigators concluded that neither output processes nor auditory temporal processing could account for the working memory deficits in children with SLI. They suggested instead that children with SLI have problems retaining or transforming phonological codes, particularly on tasks requiring multiple mental operations. They further speculated that these capacity limitations in working memory may be due to rapid decay of phonological representations or to performance limitations involving the use of less demanding coding and retrieval strategies.

In conclusion, there is considerable evidence that children with SLI have limitations in working memory, yet there are a number of unresolved issues. In light of the known heterogeneity of the SLI population, it seems unlikely that any single factor can account for the language difficulties of all children. Additional research is warranted to examine individual variation within this population and to explore whether limitations in working memory are differentially implicated in various subtypes of SLI. Another important issue pertains to whether deficits in working memory capacity are restricted to processing of verbal material or extend to nonverbal information as well. That is, it is important to determine whether the evidence supports a domain-specific model or a generalized capacity deficit model. Finally, future studies should determine whether memory limitations are actually a causal factor in SLI, an outgrowth of the language problems, or an independent area of difficulty for children with language disorder.

—Susan Ellis Weismer

References


**Further Readings**


**Mental Retardation**

Mental retardation is characterized by “significantly subaverage intellectual functioning, existing concomitantly with related limitations in two or more of the following adaptive skills areas: communication, self-care, home-living, social skills, community use, self-direction, health and safety, functional academics, leisure and work” (Luckerson et al., 1992, p. 5). Mental retardation thus applies to a broad range of children and adults, from those with mild deficits who function fairly well in society to those with extremely severe deficits who require a range of support in order to function. Regardless of the extent of mental retardation, the likelihood that communication development will be delayed is high. In fact, language delays or disorders are often an early outward signal of mental retardation.

Prior to the 1960s, a child who was diagnosed with mental retardation received little or no attention from investigators or practitioners in communication disorders because it was thought that the child could not learn and thus would make few gains in speech development. Following changes in policy and federal legislation, the 1960s saw the emergence of the modern scientific study of mental retardation. Since then, significant research findings about language and communication development have enhanced the speech, language, and communication outcomes for children and adults with mental retardation.

With respect to communication, children and adults with mental retardation can be broadly divided according to whether or not the individual speaks. Most children and adults with mental retardation or developmental disabilities do learn to communicate through speech, either spontaneously or with the aid of speech...
and language intervention during the developmental period (Rosenberg and Abbeduto, 1993). A substantial body of research has addressed the language and communication abilities of children and adults with mental retardation who speak. In particular, strong empirical findings about the communication abilities of children and adults with Down syndrome, fragile X syndrome, and Williams syndrome suggest a complex picture, with different relations between language comprehension and production and between language and cognition. The development of communicative and language intervention approaches for children with mental retardation who speak is an area of remarkable developments (Kaiser, 1993). Psycholinguistic research findings and behavioral instructional procedures have provided the foundation for language intervention protocols for teaching children with mental retardation specific speech and language skills. An early emphasis on direct instruction was followed by a shift away from the formal aspects of language and toward the teaching of lexical and pragmatic skills, measuring generalization, and the use of intervention approaches in a natural environment to promote the child’s social competence. Techniques include milieu teaching, parent-implemented intervention, and peer-mediated approaches. These are each identifiable, distinct language interventions with supporting empirical evidence that they work. Perhaps the most important recent development is the extension of intervention approaches to infants and toddlers with developmental disabilities, a move reflecting examinations of interventions targeted to intentional communication and language comprehension (Bricker, 1993). Overall, the field has developed by expanding the content and focus of intervention programs and fine-tuning the procedures used to deliver the interventions. Greater sophistication in language intervention strategies now permits an examination of the relationship between the characteristics a child brings to the intervention and the attributes of the intervention itself.

Some children and adults with mental retardation, however, encounter significant difficulty developing oral communication skills. Such difficulty during childhood results in inability to express oneself, to maintain social contact with family, to develop friendships, and to function successfully in school. As the child moves through adolescence and into adulthood, inability to communicate continues to compromise his or her ability to participate in society, from accessing education and employment to engaging in leisure activities and personal relationships. For the most part, individuals who experience considerable difficulty communicating are those with the most significant degrees of mental retardation. They may also exhibit other disabilities, including seizure disorders, cerebral palsy, sensory impairments, or maladaptive behaviors. They range in age from very young children just beginning development to adults with a broad range of life experiences, including a history of institutionalization. These children and adults can and now do benefit from language and communication intervention focusing on the development and use of functional communication skills, although the areas of concentration vary with age and experience.

One intervention approach that has been developed for use with individuals with severe communication difficulties is augmentative and alternative communication (AAC). AAC encompasses all forms of communication, from simple gestures, manual signs and picture communication boards to American Sign Language and sophisticated computer-based devices that can speak in phrases and sentences for their users. Children with mental retardation who can benefit from AAC are usually identified based on communication profiles. The majority of children with mental retardation who use AAC have more severe forms of mental retardation. These children never develop any speech, or develop only a few words, or are echolalic. For them, AAC provides a means with which to develop receptive and expressive language skills (Romski and Sevcik, 1996).

See also COMMUNICATION SKILLS OF PEOPLE WITH DOWN SYNDROME; MENTAL RETARDATION AND SPEECH IN CHILDREN.

—Mary Ann Romski and Rose A. Sevcik

References


Further Readings


computation of larger structures from lexical atoms. Concerning the lexicon versus issues concerning the as suggested by recent generative theory, into issues that characterize observed errors. The discussion is divided, natural classes that linguistic theory motivates or to unify superficially diverse markers of a disorder into theories of disorders. The task of such explanations is of grammatically based explanations are available to linguistic theory. The goal is to illustrate what kinds disorders. It presupposes familiarity with generative theories of disorders. It presupposes familiarity with generative linguistic theory. The goal is to illustrate what kinds of grammatically based explanations are available to theories of disorders. The task of such explanations is to unify superficially diverse markers of a disorder into natural classes that linguistic theory motivates or to characterize observed errors. The discussion is divided, as suggested by recent generative theory, into issues concerning the lexicon versus issues concerning the computation of larger structures from lexical atoms.

The lexicon can be subdivided along various dimensions, any of which might be relevant in capturing dissociations found in nonnormal language development. A primary split divides content words, i.e., nouns, verbs, adjectives, and adverbs, from all other morphemes, i.e., functional or closed-class elements. (I ignore the challenging issue of where adpositions belong; they likely are heterogeneous, a dichotomy among them being evinced in adult disorders [Rizzi, 1985; Grodzinsky, 1990].) In Chomskyan syntax of the 1980s (cf. Emonds, 1985; Chomsky, 1986; Fukui, 1986) there was a seemingly arbitrary split between functional meanings that were treated as autonomous syntactic functional categories (e.g., Tense, Determiner, Agreement, Complementizer) and those that were not (e.g., number marking on nouns, participial affixes, infinitival suffixes on verbs). Some attempts were made to understand disorders of acquisition in terms of this heterogeneous system (cf. Leonard, 1998). However, the harder linguists have looked, the more functional heads for which they have found structural evidence. On parsimony grounds we now expect all functional meanings to be represented in syntactic positions separate from those of content words (cf. van Gelderen, 1993; Hoekstra, 2000; Jakubowicz and Nash, 2001).

Among functional elements (morphemes or features thereof) a further distinction is made, dubbed “interpretability” by Chomsky (1995). Some morphemes encode elements of meaning and contribute directly to the interpretation of a sentence; for example, the English noun plural suffix -s combines with a noun whose meaning describes a kind of entity, and adds information about number; for example, dog + s = cani-nehood + more-than-one. Plural -s is therefore an interpretable morpheme. This contrasts with the 3sg present indicative verbal suffix -s, which makes no semantic contribution, because what is semantically “one” or “more than one” is the subject of the sentence—a noun phrase. Numerosity (and person) are properties of noun meanings, not verb meanings. The inflection -s appears on the verb as a consequence of the number and person of the subject, that information having been copied onto the verb. Agreement is therefore an uninterpretable morpheme: it replicates a bit of meaning represented elsewhere, surfacing only because the morphosyntax of English requires it. It has no impact at Logical Form. Such morphemes are seen as a plausible locus for impairment (cf. Clahsen, Bartke, and Göllner, 1997).

A second common kind of uninterpretable morphology is case marking. Many case markings have nothing to do with meaning. For example, in She saw me and I was seen by her, the difference between accusative me and nominative I does not correspond to any change in the semantic role played by the speaker; rather, it reflects purely syntactic information. A third kind of uninterpretable morphology arises in concord. For instance, in Latin the forms of many words within a noun phrase reflect its case and number, as well as the gender of the head noun: for example, ill-as vi-as angust-as “those-acc.fem.pl streets-acc.fem.pl narrow-acc.fem.pl."

Among the uninterpretable features there may be a further distinction to be drawn, as follows. Person and number information is interpretable on noun phrases but not on verbs; in a sense, there is an asymmetry between the contentful versus the merely duplicated instantiation of those features (cf. Clahsen, 1989). Case is different: it is taken to be uninterpretable both on the recipient/checkee (noun phrase) and on the source/checker (verb, preposition, or Tense); it does not exist at Logical Form at all. Thus, symmetrically uninterpretable functional elements such as Case constitute another natural class. Grammatical gender is also an example of a morphosyntactic feature that has no semantic counterpart. There are two ways in which gender marking might be impaired, calling for different explanations: first, children might not be able to consistently recall the gender of particular nouns; second, children might not be able to consistently copy gender information from the noun, which is associated with gender in the lexicon, to other parts of a noun phrase (a problem with concord).

Turning now from the lexicon to the syntax, under Minimalism the transformational operations Merge and Move are restricted in a way that effectively builds some former filters into their definitions. Taking this seriously, certain ways of talking about language disorders in terms of missing or dysfunctional subcomponents of syntax no longer make much sense. For example, in the theory of Chomsky (1981), it was plausible to talk about, say, bounding theory (which encompasses the constraints on movement operations) being inoperative as a consequence of some disorder; this simply meant that certain structures that were always generable by Move a were no longer declared invalid after the fact by virtue of a movement that was too long. There is no natural translation of this idea into Minimalist machin-
ery because locality is part of what defines an operation as a Move. A language disorder could in principle involve a change in this definition, but we cannot think of this as simply excising a piece of the grammar. Similarly, since movement in Minimalism is just a means to an end, it makes little sense to think of eliminating movement while leaving the rest of the theory intact; the system is now designed such that movement is the only way to satisfy a fundamental requirement that drives the computational system, namely, the need to create valid Logical Forms by eliminating uninterpretable features.

The syntactic accounts of language disorders that have been suggested often involve impoverished structures. One way to execute this is to posit that particular functional heads are either missing from syntactic structures or do not contain (all) the features that they would for adults (cf. Waxler, 1998). Another way is by reference to the position of the heads in the clausal structure rather than by reference to their content. Thus, it can be proposed that all structure above a particular point, say VP, is either missing (Radford, 1990) or optionally truncated (Rizzi, 1994). The extent to which particularly the tree pruning variant of this idea (Friedmann and Grodzinsky, 1997) can be characterized as simply lesioning one independently motivated piece of grammar is open to debate.

The division between stored representations and computational procedures is also relevant in (particularly inflectional) morphology, where the contrast between general rules and stored probabilistic associations has a long history (Pinker, 1999). Where precisely the line should be drawn differs, depending on one’s theory of morphology. For instance, if there are any rules at all, then surely the process that adds (a predictable allo morph of) -d to form the past tense of a verb is a rule, and unless all morphology is seen as executed by rules, the relationship between am and was in the lexicon as an idiosyncratic fact about the verb be, not encoded as a rule. But in between lie numerous subregularities that could be treated either way. For example, the alternation in sing-sang, drink-drank, and so on could be represented as a family of memorized associations or as a rule (with some memorized exceptions) triggered by the phonological shape of the stem ([ŋ]). Most psychologically oriented research has assumed that there is only one true rule for each inflectional feature such as [past tense] (supplying the default), while the generative phonology tradition has used numerous rules to capture subpatterns within a paradigm, only the most general of which corresponds to the default in having no explicit restriction on its domain. Thus, although a dissociation between rules and stored forms is expected under virtually any approach, precise predictions vary.

There are several ways in which (inflectional) morphology might be impaired. One of the two mechanisms might be entirely inopervative, in which case either every inflected form would have to be memorized by rote (no rules; cf. Gopnik, 1994) or every word would have to be treated as regular (no associations). The latter would be evinced by overregularizations. The former might yield overirregularizations if, in the absence of a rule, analogy-based associations were unchecked and overextended; alternatively, it might have a subtler symptomology whereby any verb could be correctly learned once its past tense was heard, but in a Wug-testing situation nonce forms could not be inflected. A more moderate impairment might entail that at least some irregular forms would be demonstrably learnable, but in production they would not always be retrieved reliably or quickly enough to block a general rule (Clahsen and Temple, 2002), apparently violating the Elsewhere Condition.

We turn now to the logic governing ways in which the most commonly produced error types from the child language disorder literature can be explained. Under the strong separation of computational combinatoric machinery versus lexical storage pursued in Minimalist syntax, it is possible for children with normal syntactic structures to sound very unlike adults, because in their lexicon certain morphemes either are missing or have incorrect features associated with them. For example, the fact that some children learning English might never produce -s in an obligatory 3sg context could be consistent with them having mastered the syntax of agreement, if their lexicon has a missing or incorrect entry for 3sg. Consequently, it is important to document the lexical inventory, that is, whether each child does at least sometimes produce the forms of interest or can in some way be shown to know them. Similarly, suppose that children learning a given language produce agreement mismatch errors of just one type, namely, that non-3sg subjects appear with 3sg verb forms. That is, suppose that in Spanish we find errors like yo tiene (“I have (3sg)”) but no errors like ella tengo (“she have (1sg)”). One could postulate that 3sg forms (e.g., tiene) are unspecified for person and number features and are inserted whenever more specific finite forms (e.g., tengo) are unavailable, or that tiene has wrongly been learned as a 1sg form. In either case, again, the syntax might be fully intact.

Errors due to lexical gaps should pattern differently from errors due to syntactically absent heads. Consider the Tense head as an example. The meaning of tense itself (past versus present) can be expressed morphologically in the Tense head position, and in addition Tense is commonly thought to house the uninterpretable feature that requires the presence of an overt subject in non-null subject languages (the unhelpfully named EPP feature), and the uninterpretable feature that licenses nominative case on its specifier (the subject). If Tense were completely missing from the grammar of some children, this would predict not only that they would produce no tense morphemes, but also that they would not enforce the overt subject requirement and would not (syntactically) require nominative case on subjects. If, on the other hand, what is observed is not an absence of tense marking but rather an incorrect choice of how to express Tense morphophonologically (e.g., singed, instead of sang), this would not be compatible with absence of the Tense head and would not predict the other syntactic consequences mentioned. Only omission of an inflectional marking or perhaps supernumerary inflection
(e.g., Did he cried?) could have a syntactic cause; incorrect allomorph selection could not.

If a feature such as Tense is expressed some of the time but not always, this could have two underlying causes that lead to different syntactic expectations. One possibility is that Tense is part of the syntactic representation in all cases and its inconsistent expression reflects some intermittent problem in its morphophonological spell-out (cf. Phillips, 1995; Marcus, 1995); in that case no syntactic consequences are predicted. The other possibility is that Tense is intermittently absent from syntactic representations. In just this scenario we expect that the syntactic properties controlled by the Tense head should be variable (e.g., subjects are sometimes nominative, sometimes not), and furthermore, utterance by utterance, syntax and morphology should correlate: the Tense morpheme should be missing if and only if nominative case is not assigned/checked and the overt subject requirement is not enforced.

It is crucial to understand such predicted contingencies as claims concerning the distribution of contrasting forms. It seems clear that some children go through a stage during which, for example, some of the English pronoun forms are not produced at all; this is particularly common for she. What behavior we should expect at this stage depends on assumptions about the architecture of the syntax-morphology interface. Taking an “early insertion” view, under which only complete words from the lexicon can be inserted in syntactic derivations, a child lacking a lexical entry for the features [pron, 3sg, fem, NOM] should be unable to generate a sentence requiring such an item. In contrast, on a “late insertion” nonblocking approach such as Distributed Morphology (Halle and Marantz, 1993), there is always some form (perhaps phonologically null) that can be used to realize the output of a syntactically valid derivation. The syntactic tree is built up using feature bundles such as [pron, 3sg, fem, NOM], without regard to which vocabulary item might fill such a position. The architecture dictates that if there is no vocabulary entry with exactly this set of features, then the item containing the greatest subset of these features will be inserted. Thus, a child who knows the word her and knows that it is a feminine singular pronoun could insert it in such a tree, producing Her goes from a fully adultlike finite clause structure.

We conclude with some methodological points. Generative grammar’s conception of linguistic knowledge as a cognitive (hence also neural) representation in the mind of an individual at some particular time dictates that, in analyzing behavioral data collected over an extended period or from multiple children, pooled data cannot be directly interpreted at face value. For example, suppose we analyze transcripts of three children’s spontaneous productions sampled over a period of 6 months and find that in obligatory contexts for some grammatical morpheme, say 3sg -s, the overall rate of use is 67%. Virtually nothing can be concluded from this datum. It could represent (among other possibilities) a scenario in which two children were producing -s at 100% (i.e., talking like adults in this regard) and the third child never produced -s. In this circumstance we would have no evidence of a developmental stage at which -s is optional, and the third child’s productions would be consistent with her simply not knowing the form of the 3sg present tense verbal inflection in English. At another extreme, if each of the three children were producing -s in two-thirds of obligatory contexts, we would have to posit a grammar in which Tense or Agreement are optional (or else multiple concurrent grammars). The same logic applies in the temporal dimension: a period of 0% production followed by a period of 100% production calls for a different analysis from a period during which production is at 50% within single recording sessions. Therefore, data reporting needs to facilitate assessment of the extent to which samples being pooled represent qualitatively comparable grammars. In addition to measures of central tendency and variance, this calls for the kind of distributional information found in “box-and-whiskers” plots.

Furthermore, to say that a child “has acquired X” by a certain age, where X is some morpheme or class of morphemes, is not strictly meaningful from the linguistic perspective adopted here. We can speak of attaining levels of normal or adult performance, and we can speak of having established a lexical entry with the correct feature specifications, but there is nothing in grammar that could correspond to a claim such as the following: “A child is taken to have acquired X once her rate of production of X morphemes in obligatory contexts is greater than 90%” (cf. Brown, 1973). This begs the question of what the grammar was like when production of X was at 85%: If X was not at that time a part of her grammar, how did the child manage to create the illusion of using it correctly so much of the time? Also, rates of production in obligatory contexts must be complemented by correct usage rates; one is scarcely interpretable without the other. If X is always used correctly, then even very low production rates signal knowledge of the properties of X and the syntactic conditions on its distribution. But if X is also frequently used incorrectly, neither type of knowledge can be inferred.

——Carson T. Schütze

References


Further Readings


During the first few years of life, a time that is critical for language development, a relationship between otitis media with effusion (OME) and associated hearing loss in early childhood has been hypothesized to affect children’s language acquisition in the areas of phonology, vocabulary, syntax, and discourse in several ways. First, the disruption and variability in auditory input due to OME may cause children to encode information incompletely and inaccurately into their phonological working memory. Consequently, children’s lexical development may be hindered if they have inaccurate representations of words, which may then result in imprecise lexical recognition or production. Second, OME-associated hearing loss may result in difficulties acquiring inflectional morphology and grammar. Children may not hear or may inaccurately hear certain grammatical morphemes that are of low phonetic substance, such as inflections of short duration and low intensity (e.g., third person /s/, past tense /’ed’/) and unstressed function words (“is,” “the”). Third, children’s use of language may also be affected because they may miss subtle nuances of language (e.g., intonation marks, questions), which interferes with their ability to follow conversations. Children with prolonged or frequent OME may also learn to tune out, particularly in noisy situations, resulting in attention difficulties for auditory-based information. Difficulty maintaining sustained attention could compromise children’s ability to sustain discourse (i.e., to follow and elaborate on the topic of the conversation) and to organize and produce coherent narratives (both requiring auditory memory and recall).

Recent models of a potential linkage between a history of OME and subsequent impaired language development hypothesize that not only factors inherent in the child but also the child’s environment and the interaction between the child and the environment can affect this relationship (Roberts and Wallace, 1997; Vernon-Feagans, Emanuel, and Blood, 1997; Roberts et al., 1998; Vernon-Feagans, 1999). These additional factors include both risk factors (e.g., the child has poor phonemic awareness skills, the mother has less than a high school education, the child care environment is noisy) and protective factors (e.g., the child has an excellent vocabulary, a literacy-rich home environment, and a responsive child care environment). Thus, it is proposed that the potential impact of OME on children’s language development depends on the number and timing of OME episodes and associated hearing loss; the child’s cognitive, linguistic, and perceptual abilities; the responsiveness and supportiveness of the child’s environment; and interactions among these variables.

Over the past three decades, more than 90 original studies have examined whether children who had frequent episodes of OME in early childhood scored lower on measures of language than children without such a history. Earlier studies examining an association between OME and later language were retrospective in design (the children’s history of OME was documented by parents reporting the frequency with which children...
had OME or by a review of medical records collected by different medical providers) and were more likely to contain measurement errors. More recent studies of the OME–language linkage were prospective, with children’s OME histories documented longitudinally from early infancy and repeated at specific sampling intervals. Prospective studies are more likely to have greater objectivity and accuracy over time, avoiding many of the methodological limitations of previous studies.

Several prospective studies have found a relationship between a history of otitis media in early childhood and later language skills during the preschool and early elementary school years. More specifically, in comparison with children who infrequently experienced otitis media, infants and preschoolers with a history of OME scored lower on standardized assessments of receptive and expressive language (Teele et al., 1984; Wallace et al., 1988; Friel-Patti and Finitzo-Hieber, 1990) and in specific language areas, including syntax (Teele et al., 1990), vocabulary (Teele et al., 1984), and narratives (Feagans et al., 1987). However, many studies failed to find associations between an early history of OME and later measures of overall receptive or expressive language, vocabulary, or syntax (Teele et al., 1990; Peters et al., 1994; Paradise et al., 2000; Roberts et al., 2000).

Several ongoing prospective studies are providing new and important information on whether a history of OME in early childhood causes later language difficulties. Three recent experimental studies (Maw et al., 1999; Rovers et al., 2000; Paradise et al., 2001) examined whether prompt insertion of tympanostomy tubes (to drain the fluid for children with frequent or persistent OME) improved children’s language development, compared with delaying the insertion of tympanostomy tubes. Paradise and colleagues (2001) randomized 429 children (at mean age of 15 months) who had persistent or frequent OME to have tympanostomy tubes inserted either promptly or 6–9 months later and reported no language differences between the two treatment groups at age 3 years of age. Rovers and colleagues (2000) also did not find that prompt insertion of tympanostomy tubes improved children’s language development. Maw and colleagues (1999) did find effects on language development 9 months after treatment; however, 18 months after treatment there were no longer differences between the groups.

Other prospective studies considered the impact of multiple factors such as the educational level of the mother and the extent of hearing loss a child experienced during early childhood on children’s language development. The Pittsburgh group (Feldman et al., 1999; Paradise et al., 2000) reported weak but significant correlations between OME in the first 3 years of life and language development (accounting for 1%-3% of the variance in language skills), after controlling for many family background variables. Roberts and colleagues (1995, 1998, 2000) prospectively studied the relationship of both children’s OME and hearing history to language development. They did not find a direct relationship between OME or hearing history and children’s language skills between 1 and 5 years of age (Roberts et al., 1995, 1998, 2000). They did find that the caregiving environment (responsiveness of the child’s home and child care environments) mediated the relationship between children’s history of OME and associated hearing loss and later communication development at 1 and 2 years of age (Roberts et al., 1995, 1998, 2000, 2002). That is, children with more OME and associated hearing loss tended to live in less responsive caregiving environments, and these environments were linked to lower performance on measures of receptive and expressive language skills. More recently, Roberts and colleagues reported that children with greater incidence of OME scored lower in expressive language upon entering school but caught up with their peers in expressive language by second grade. However, a child’s home environment was much more strongly related to early expressive language skills than was OME. These and other ongoing prospective studies highlight the importance of examining the multiple factors that affect children’s language development.

The potential impact of frequent and persistent hearing loss due to OME on later language skills may be particularly important to examine in children from special populations who are already at risk for language and learning difficulties. Children who have Down syndrome, fragile X syndrome, Turner’s syndrome, Williams’s syndrome, cleft palate, and other craniofacial differences often experience frequent and persistent OME in early childhood (Zeisel and Roberts, 2003; Casselbrant and Mandel, 1999). This increased risk for OME among special populations may be due to craniofacial structural abnormalities, hypotonia, or immune system deficiencies. A few retrospective studies have reported that a history of OME further delays the language development of children from special populations (Whiteman, Simpson, and Compton, 1986; Lonigan et al., 1992).

The question of whether recurrent OME affects the later acquisition of language is still unresolved, in part because of the conflicting findings of studies that have examined this issue. There is increasing support from prospective studies that for typically developing children, OME may not be a substantial risk factor for later language development. Although a few studies report a very mild association between OME and later impairment of receptive and expressive language skills during infancy and the preschool years, the effect is generally very small, accounting for only about 1%-4% of the variance. Furthermore, it is clear that the caregiving environment at home and in child care plays a much more important role than OME in children’s language development. Future research should examine if frequent hearing loss due to OME relates to children’s language development. The impact of a history of OME and associated hearing loss on the language development of children from special populations should also be further studied. Some typically developing children as well as children from special populations may be at increased risk for later language and learning difficulties due to a
history of OME and associated hearing loss. Until further research can resolve whether such a relationship between a chronic history of OME and later language skills exists and can determine what aspects of language are affected, hearing status and language skills need to be considered in the management of young children with histories of OME.

Several strategies have been recommended for young children who are experiencing chronic OME (Roberts and Medley, 1995; Roberts and Wallace, 1997; Vernon-Feagans, 1999; Roberts and Zeisel, 2000). First, a child's hearing, speech, and language should be tested after 3 months of bilateral OME, or after four to six episodes of otitis media in a 6-month period, or when families or caregivers are concerned about a child's development. Second, families and other caregivers (e.g., child care providers) of young children with recurrent or persistent OME need clear and accurate information in order to make decisions about the child's medical and educational management. Third, children who experience recurrent or persistent OME, similar to all children, will benefit from a highly responsive language- and literacy-enriched environment. Caregivers should respond to communication attempts, provide frequent opportunities for children to participate in conversations, and read often to their children. Fourth, children with chronic OME will benefit from an optimal listening environment in which the speech signal is easy to hear and background noise is kept to a minimum. Fifth, some children with a history of OME may exhibit language and other developmental difficulties, and benefit from early intervention. Finally, the results of ongoing research studies combined with previous studies should help determine whether a history of OME in early childhood places children at risk for later language difficulties, and if so, how to then target intervention strategies.

—Joanne E. Roberts

References


Further Readings


Perseveration

Perseveration, a term introduced by Neisser in 1895, refers to the inappropriate continuation or repetition of an earlier response after a change in task requirements. Although individuals without brain damage may display occasional perseverative behaviors (e.g., Ramage et al., 1999), as Allison (1966) pointed out, when perseveration is pronounced, “it is a reliable, if not a pathognomonic sign of disturbed brain function” (p. 1029). Indeed, perseveration has been described in association with a variety of neurological and psychiatric conditions, including stroke, head injury, dementia, Parkinson’s disease, and schizophrenia.

Perseveration is such a notable and fascinating clinical phenomenon that for more than 100 years, various researchers have attempted to more precisely describe its characteristics, label its subtypes, and identify its neuropsychological correlates and neuropsychological mechanisms. Good agreement has emerged from these studies as to the characteristics of various forms of perseveration, with rather little agreement as to labels for subtypes. And although most investigators agree that the frontal lobes and their associated white matter pathways play a prominent role in perseveration, other areas of the brain have been implicated. The neuropsychological mechanisms responsible for perseveration also are uncertain and probably vary according to subtypes. Among the mechanisms implicated are persistent memory traces, failure to inhibit prepotent responses, pathological inertia, and failure to disengage attention. For a review of some of this literature, see Hotz and Helm-Estabrooks (1995a).

It is important for professionals working with individuals having neurological conditions to be aware of perseveration and recognize its subtypes, because this behavior can contaminate experimental and clinical test results and reduce communicative effectiveness. Perseveration can occur in any behavioral output modality, including speech, writing, gesturing, drawing, and other forms of construction. Three primary forms of perseveration have been described, with one of these forms having four subtypes. The terms used here are derived from several sources (e.g., Santo-Pietro and Rigrodsky, 1986; Sandson and Albert, 1987; Lundgren et al., 1994; Hotz and Helm-Estabrooks, 1995b).

Stuck-in-set perseveration is the inappropriate maintenance of a category or framework of response after introduction of a new task. For example, as a part of a standardized test, an individual with traumatic brain injury without aphasia was asked to list as many animals as he could in 1 minute. He listed ten animals before he was given the following instructions: “Now I want you to name as many words as you can that start with the letter M. [He was shown a lowercase m.] Here are the rules. Do not name words that begin with a capital M. Do not say the same words with a different ending, like mop, then mopped or mopping. [The written letter m was removed.] Okay, you have 1 minute to name as many
words you can think of that start with the letter m.” In response, the man said “monkey,” “mouse” in the first few seconds, then “man” after 15 seconds. He produced no further responses for the remaining time. Thus, although he understood the concept of listing m words, he could not disengage from the idea of listing animals, and his score for producing words according to a letter/sound category was contaminated by stuck-in-set perseveration.

**Continuous perseveration** is the inappropriate prolongation or continuation of a behavior without an intervening response or stimulus. For example, a woman with Alzheimer’s disease was given the following spoken and written instructions: “Draw a clock. Put in all the numbers. Set the hands to 10 minutes after 11.” She wrote numbers 1 through 18 in the circle provided before she ran out of space. She then drew a hand to the number 10, but continued to draw hands to each number. Thus, either she was unable to disengage from the idea of drawing clock hands or she was unable to inhibit that particular graphomotor activity.

**Recurrent perseveration** is the inappropriate recurrence of a previous response following presentation of a new stimulus or after giving a different intervening response. For example, a man with fluent aphasia was asked to write the days of the week. He wrote, “Monday, Tuesday, Wednesday, Tuesday, Friday, Saturday, Monday, Sunday.”

Various subtypes of recurrent perseveration have been described and labeled. A primary distinction can be made between carryover of part of the phonemic structure of a previous word and repetition of an entire word.

An example of **phonemic carryover perseveration**, in which part of the phonemic makeup of a previous word is inappropriately repeated, is “comb” for comb, then “klower” for flower.

Within the category of **whole-word carryover**, three types of perseveration occur, semantic, lexical, and program-of-action perseveration.

**Semantic perseverations** are words that are semantically related to the target (e.g., repetition of the naming response apple when shown a lemon).

**Lexical perseverations** are words that have no obvious semantic relation to the target (e.g., repetition of the word key when asked to name scissors).

**Program-of-action perseverations** are repeated words that begin with the same initial sound as a previous response (e.g., repetition of the response wristwatch for subsequent objects, such as wrench, whose names begin with /w/).

Although, as mentioned earlier, perseveration occurs in association with many neurological and psychiatric conditions, perseverative behavior is especially notable in acquired aphasia. The results of their study of perseverative behaviors in aphasic individuals prompted Albert and Sandson (1986) to suggest that it “may even comprise an integral part of [the] specific language deficits in aphasia” (p. 105). This suggestion is supported by the work of other investigators (e.g., Santo-Pietro and Rigodsky, 1982; Emery and Helm-Estabrooks, 1989; Helm-Estabrooks et al., 1998). There is good evidence that perseverative behaviors are unrelated to time post onset of aphasia but are correlated significantly with aphasia severity. Thus, perseveration can be a persistent problem for individuals with aphasia and interfere with all modalities of communicative expression. As such, perseveration is an important treatment target for speech and language clinicians working with aphasic individuals, although few approaches have been described thus far. Exceptions are the program designed by Helm-Estabrooks, Emery, and Albert (1987) to reduce verbal recurrent perseveration, and the strategies described by Bryant, Emery, and Helm-Estabrooks (1994) to manage various forms of perseveration in severe aphasia. See Helm-Estabrooks and Albert (2003) for updated descriptions on these methods.

—Nancy Helm-Estabrooks

**References**


Further Readings


Phonological Analysis of Language Disorders in Aphasia

Various approaches have been utilized in the study of phonological disorders in aphasia, but each usually shares certain assumptions of the others. This overview presents these different orientations with the recognition that there is a good deal of overlap in each.

Neurolinguistic

Renee Beland (and several colleagues) have used “underspecification theory” to capture the types of phonological errors in fluent aphasics and the constraints on those errors (Beland, 1990; Beland, Caplan, and Nespoulos, 1990). The phonological model here has three levels: the minimally specified level, a lexical level, and a surface level. Feature markedness, phonotactic patterns, and syllable constituent slots (onset, rime, nucleus, coda) are then used to describe the nature and location of phonemic paraphasias and the constraints on those paraphasias.

As a student of Roman Jakobson, Sheila Blumstein (1973) is the intermediary between Prague School phonological studies of aphasic errors (e.g., Jakobson, 1968) and a host of neurolinguistic studies of paraphasia published since the early 1970s. Her contributions to the study of the neuropsychology and neurobiology of human language sound structure are herculean (see Blumstein, 1995, 1998, for discussion of much of her work). Her initial study of the typology of phonemic paraphasia set the stage for numerous ensuing studies.

Susan Kohn (1989, 1993) has provided a wealth of information on phonological breakdown in fluent aphasics, usually patients with conduction aphasia and Wernicke’s aphasia. Her studies have focused on the difficulties these patients have with constructing phonemic strings once the full-form lexical representations have been accessed in their stored form. Through the mechanisms of models (Shattuck-Hufnagel, 1979; Garrett, 1984), Kohn has successfully characterized many paraphasic types, located the production process where the error occurs, and specified which type of error is diagnostic of distinct aphasic syndromes. She has incorporated syllable structure constraints, phonotactic patterns, and the principle of sonority. Furthermore, Kohn, Smith, and Alexander (1996) have charted the recovery patterns of Wernicke’s patients who in the acute stages produced neologistic jargon. They observed that in certain of these patients the aphasia resolved to a chronic stage in which the patients were clearly getting closer to underlying lexical representations and producing less severe phonemic paraphasia, with lower number of random segments in those errors. In another set of patients the aphasia resolved to a chronic stage in which the patients maintained severe lexical access disruptions, producing less neology but more lexical blocks, circumlocutions, and other errors indicative of lingering anomia.

David Caplan (1987) and colleagues (Caplan and Waters, 1992) have published widely on phonological breakdown in aphasia. He, as well as Kohn, has analyzed the phonemic string construction difficulties of the syndrome of “reproduction” conduction aphasia. He localizes this disruption at the point where final sound production is called for from either a semantic input, an auditory/lexical input, or a written input. That is, the patient will produce phonemic errors in object naming, in repetition, or in reading aloud. Any one patient may produce paraphasias with all or certain ones of these inputs. Cognitive neuropsychological dissociations are numerous here.

Buckingham and Kertesz (1976) analyzed the neologistic jargon of several patients with fluent aphasia. Each patient revealed a good deal of phonemic errors, but many errors rendered underlying targets opaque. Other forms, however, were opaque and otherwise abstruse where there was no clear transparency of any certain underlying lexical target. All patients had severe anomia, which led to the suggestion that there could be two separate productive mechanisms or processes that might give rise to the production of these non-recognizable word-form errors, but in that study the question was left unexplored. The issue was broached again in Buckingham (1977), but not until Butterworth (1979) was a specific idea put forth, that of a “random generator.” Subsequently, Buckingham (1990a) attempted to mollify the critics of the random generator by showing that the idea was in a sense a metaphor for the use and appreciation of general phonological knowledge, which for all speakers underlies the ability to recognize “possible words” in a language and to produce “nonce” forms when necessary. The issue of a dual route for the neologism has reappeared recently. Gagnon and
Schwartz (1996) found no bimodal distribution in their corpus of phonemic paraphasias and neologisms, and thus argued for a single route for their productions. In contradistinction, Kohn, Smith, and Alexander (1996) observed patients who demonstrated both routes, one for phonemic paraphasia and another that involved some sort of a “backup mechanism for reconstructing” a phonological representation when either partial or no stored phonological information about a word is made available to the production system” (p. 132), which reduces to some kind of generating device.

In Buckingham (1990b), the principle of sonority (see Ohala, 1992, for some criticism of this principle) was invoked to provide an account of the structural constraints on phonemic paraphasia, and shortly thereafter Christman (1994) extended the utilization of sonority to capture the pattern constraints on a large corpus of neologisms, where she statistically demonstrated that neologisms abided by both sonority and phonotactic dictates. A major portion of her contribution was to show how sonority could go beyond mere phonotactics to characterize neologistic word production.

**Slips, Paraphasia, and the Continuity Thesis**

Slips-of-the-tongue have always played a role in the modeling, characterization, analysis, and explanation of paraphasia. There is a direct line from Hughlings-Jackson in the second half of the nineteenth century to modern psycholinguistic studies of the “lapsus linguae,” all of which have assumed at least some degree of continuity between the functional errors in normality and the paraphasias in pathology (Buckingham, 1999).

**Phonological Breakdown and Connectionist Modeling**

Wheeler and Touretzky (1997) have combined a model of segmental “licensing” from phonological auto-segmental theory with a connectionist simulation of phonological disruptions in fluent aphasia with no psycholinguistic mechanisms at all. In an impressive and surprising fashion, they were able to simulate most error types and conditions described in such work as Buckingham (1986). Dell et al. (1997) have published one of the most in-depth investigations of a connectionist system (of the interactive activation type) that simulated normal naming and aphasic naming—and only naming. The three-level (semantic, lexical, segmental) system had upward and downward feedforward and feedback connections. The system was set up so that connection weight values and decay rate values could be varied globally throughout the system as a whole. Phonological errors (and other errors) that indicated level interaction were produced by lesioning only the decay rates; phonemic paraphasias that did not render targets opaque and formal verbal substitutions where error and target shared phonological features. Lesioning only connection weights, however, simulated phonological errors (and other errors) that did not indicate interaction between levels:

neologisms, where target words were not recognizable. Recovery patterns were then analyzed and simulated, where it appeared to the authors that although there was improvement in productions, the productions themselves remained either of the decay lesion type or of the connection lesion type. The only problem with phonological breakdown would be that those who had produced neologisms early on may resolve to a less severe phonemic paraphasia, where targets could be increasingly gleaned, thereby indicating level interaction. That scenario, of course, would mean that a connection weight lesion pattern would resolve to a decay rate lesion pattern. Most connectionist-oriented accounts do not provide for dual routes for neologisms (e.g., Gagnon and Schwartz, 1996; Gagnon et al., 1997; Goldman, Schwartz, and Wilshire, 2001; Nadeau, 2001), locating their production exclusively between the lexeme and the phoneme strata. For instance, the model of Hillis et al. (1999, p. 1820) accounts for phonemic paraphasias and neologisms with different degrees of connectionist weakenings between these two strata, where only a “few” nontarget subword units would be activated for phonemic paraphasias but “many” nontarget subword units would be activated for neologisms. It is not yet clear just how connectionist models will continue to eschew dual routes for neologisms.

**Apraxia of Speech Meets Phonemic Paraphasia: Phonology or Phonetics?**

Square, Roy, and Martin (1997) have recently discussed posterior lesion apraxia of speech patients, which accords with ideas suggested in Buckingham and Yule (1987). In addition, many presumed phonemic paraphasias in Broca’s aphasia (e.g., Keller, 1978) could very plausibly stem from apraxic asynchronies (Buckingham and Yule, 1987). Speech perception disturbances have been observed in Broca’s aphasics, and Wernicke’s aphasics often present with articulatory abnormalities (Price, Indefrey, and van Turennout, 1999, p. 212). The perceptual function of the inferior frontal gyrus has also been observed by Hsieh et al. (2001). Galaburda has uncovered motor regions in layer III of the left temporal cortex in the plenum region (Galaburda, 1982, and Amaducci et al., 1981) have found asymmetrical concentrations of choline-acetyltransferase (ChAT) in this region from neurochemical assays at autopsy in human subjects. Recent reviews (e.g., Blumstein, 1995) have emphasized the highly distributed nature of the sound system throughout the left perisylvian cortex. Little wonder there is such a degree of indeterminacy with aphasic sound system disruptions in stroke. From purely linguistic reasoning, Ohala (1990) has challenged the claim that phonetics and phonology are separate systems; rather, they are totally integrated. Vocal tract embodiment is tightly linked to most so-called phonological properties (e.g., syllable constituency and phonotactics, sonority, feature markedness, coarticulatory processes), so that whatever element may be “phonologized” will remain forever linked to vocal tract dynamics (see
Christman, 1992; MacNeilage, 1998). It would seem that we are poised for major reevaluations in our understanding of how the human sound system breaks down in aphasia.

—Hugh W. Buckingham

References


Further Readings


Phonology and Adult Aphasia

The sound structure of language is the primary medium for language communication. As a result, deficits affecting phonology, defined as the sound structure of language, may have a critical impact on language communication in general, and specifically on the processes involved in both speaking and understanding.

Deficits in Speech Production

A number of stages of processing underlie speech production. The speaker must select a word candidate from the lexicon and access its phonological form. Once selected, the sound structure of the word or utterance must be planned—the phonological representation is encoded in terms of the phonological properties of the sound segments, their order and context, and the prosodic structure of the word as a whole. The next stage of processing is articulatory implementation, in which the more abstract phonological representation is converted into a set of motor commands or motor programs for the phonetic realization of the utterance.

There is some evidence to suggest that these stages of production may be dissociated in the aphasias (Nadeau, 2001). However, the patterns suggest that the dissociations are not complete, and hence the production system appears to be neurally distributed in the left perisylvian regions of the left hemisphere (Blumstein, 2000). In par-
ticular, patients with Broca’s aphasia and other patients with anterior lesions appear to have predominantly articulatory implementation deficits and to a lesser extent phonological selection and planning impairments. In contrast, those with Wernicke’s and conduction aphasia and patients with posterior lesions appear to have predominantly lexical selection and phonological planning deficits and to a minor degree articulatory implementation impairments.

The clearest evidence for a dissociation between the production stage of articulation implementation and the stages of phonological selection and planning comes from investigations of the acoustic patterns of speech productions. These studies show that persons with anterior aphasia, including those with Broca’s aphasia, have a deficit in the articulatory implementation of speech. Articulatory timing and laryngeal control appear to be particularly affected. The timing disorder emerges in the production of those speech sounds requiring the coordination of two independent articulators, such as the production of voicing in stop consonants, as measured by voice-onset time (Blumstein et al., 1980), and the complex timing relation between syllables (Kent and McNeil, 1987; Gandour et al., 1993). Deficits in laryngeal control are evident in the production of sound segments as well as in the production of prosody. Individuals with these deficits show lower and more variable amplitudes of glottal excitation during the production of voicing in fricative consonants (Code and Ball, 1982) as well as changes in the spectral characteristics associated with the production of stop consonants (Shinn and Blumstein, 1983). Prosodic disturbances are characterized by a restricted fundamental frequency range (Danly and Shapiro, 1982). Additionally, the production of tone in languages such as Thai and Chinese is affected, although the global properties of tone, such as whether the tone is rising or falling, are maintained (Gandour et al., 1992).

Individuals with posterior aphasia, including those with Wernicke’s and conduction aphasia, also appear to have a subtle articulatory implementation deficit that is different from that of individuals with anterior aphasia. Characteristics of this deficit include increased variability in a number of phonetic parameters, including vowel formant frequencies and vowel duration, and abnormal temporal patterns between syllables (Baum et al., 1990; Vijayan and Gandour, 1995).

Deficits in selection and planning emerge across a wide spectrum of aphasia, including left hemisphere anterior (Broca’s) and posterior (conduction and Wernicke’s) aphasia. Evidence in support of this is provided by similar patterns of phonological errors produced by these patients in spontaneous speech (Blumstein, 1973; Holloman and Drummond, 1991). Error types include phoneme substitution errors (e.g., “teams” → [kimz]), simplification errors (e.g., “brown” → [bawn]), addition errors (e.g., “bet” → [blet]), and environment errors (e.g., “degree” → [gardri] or “moon” → [mum]). These utterances deviate phonologically from the target word but are implemented correctly by the articulatory system, reflecting an inability to correctly encode or activate the correct phonemic (i.e., phonetic feature) representation of the word. Evidence that the phonological representations are intact comes from the variability with which errors occur. An affected individual may make one or a series of phonological errors on a target word and also produce it correctly (Butterworth, 1992).

Despite these similarities, differences have emerged in the patterns of production in naming and repetition tasks, suggesting a difference in the locus of the underlying impairment. Persons with conduction aphasia produce many more phonological errors than persons with Wernicke’s aphasia, and the patterns of production in individuals with conduction aphasia more nearly approximate the phonological representation of the target word than those of persons with Wernicke’s aphasia. These results suggest that the basis of the deficit in Wernicke’s aphasia more likely resides in the processes involved in lexical selection, whereas the basis of the deficit in conduction aphasia more likely lies in the processes involved in phonological planning (Kohn, 1993).

In sum, the evidence suggests that the neural system underlying speech production is a distributed neural network with functional subcomponents. Wernicke’s area and the association cortices around it are implicated in the processes of lexical selection. The supramarginal gyrus and the white matter deep to it appear to be involved in phonological selection and planning. The motor areas including the frontal operculum, the premotor and motor regions posterior and superior to the frontal operculum, and the white matter below, including the basal ganglia and insula, appear to be involved in the articulatory implementation of speech (Dronkers, 1997; Damasio, 1998). Nonetheless, the evidence suggests that there is not a 1:1 relationship between these neurological landmarks and the stages of output, since all stages of speech production appear to be affected in all of the patients, although to varying degrees.

**Deficits in Speech Perception**

A number of stages of processing have been identified in auditory word reception. These stages include the extraction of generalized auditory patterns from the acoustic waveform, the conversion of this spectral representation to a more abstract phonetic feature/phonological representation, and the mapping of this phonological representation onto lexical form (i.e., a word in the lexicon) (Nadeau, 2001). Deficits at any one or all of these levels may potentially contribute to auditory comprehension impairments.

There is little evidence to suggest that aphasic patients have deficits at the stage of extracting generalized auditory patterns from the acoustic waveform (Polster and Rose, 1998). However, they show impairments in processing a number of auditory/phonetic parameters of speech, including temporal cues such as voice-onset time, a cue to the phonetic dimension of voicing, and spectral cues such as formant transitions, cues to the phonetic dimension of place of articulation (Blumstein et al.,
1984; Shewan, Leeper, and Booth, 1984). Increasing the duration of the formant transitions to allow more time to process the rapid spectral changes associated with the perception of place of articulation does not improve the performance of aphasic patients (Riedel and Studdert-Kennedey, 1985). In general, patients have great difficulty performing these tasks, particularly when the stimuli are synthetic speech stimuli as compared to natural speech stimuli and when the task requires labeling or naming the stimuli as compared to discriminating them. Nonetheless, the patterns of performance suggest that individuals with aphasia can map the spectral representations on to the phonetic features of language. Those who can perform the labeling and discrimination tasks typically show categorical perception similar to normal subjects. They perceive the stimuli as belonging to discrete phonetic categories, and they show peaks in discrimination at the phonetic boundaries between the phonetic categories. All of the deficits described above emerge across aphasic syndromes and do not correlate with the severity of auditory comprehension impairment (Basso, Casati, and Vignolo, 1977).

In contrast to the difficulty in perceiving acoustic dimensions associated with voicing and place of articulation, persons with aphasia are generally able to perceive those acoustic dimensions contributing to the perception of speech prosody, that is, intonation and stress. Even individuals with severe auditory comprehension impairments are able to use intonation cues to recognize whether an utterance is a command, a yes–no question, an information question, or a statement (Green and Boller, 1974). Nonetheless, some impairments have emerged in the perception of more local parameters of prosody, including word stress and tone (Gandour and Dardarananda, 1983; Baum, 1998). However, no differences have emerged in the performance of persons with anterior and posterior aphasia.

Similar to studies of the acoustic-phonetic properties of speech, nearly all persons with aphasia show perceptual deficits in phonological processing (Blumstein, Baker, and Goodglass, 1977; Jauhiainen and Nuutila, 1977; Csepe et al., 2001). These deficits emerge in tasks requiring subjects to discriminate words or nonsense syllables contrasting by one or more phonetic features (e.g., “dime”—“time” or “da”—“ta”) or to point to an auditorily presented target stimulus from a visual array of objects or written nonsense syllables that are phonologically similar. Individuals with aphasia have more difficulty on labeling and pointing tasks than on discrimination tasks. They also make more errors in the perception of nonsense syllables than in the perception of real words, although the patterns of performance are similar across these stimulus types. For all patients, the perception of consonants is worse than that of vowels; more errors occur when the stimuli contrast by a single phonetic feature; and more errors occur in the perception of place of articulation and voicing than in the perception of other phonetic feature contrasts.

The results of studies investigating the perception of speech challenge the view that the posterior left hemisphere, particularly Wernicke’s area and associated temporal lobe structures, is selectively involved in speech receptive functions. Instead, the results are consistent with the view that phonetic/phonological processing has a distributed neural basis, one that involves the perisylvian regions of the left hemisphere (Hickok and Poeppel, 2000; Burton, 2001). Although persons with Wernicke’s aphasia make a large number of errors on speech perception tasks, and although damage to the left supramarginal gyrus and the bordering parietal operculum is correlated with speech perception deficits (Gow and Caplan, 1996), some persons with anterior aphasia have shown poorer performance on these tasks than those with Wernicke’s aphasia. Moreover, performance on speech perception tasks has failed to show a strong correlation with severity of auditory language comprehension. Thus, although phonetic and phonological processing deficits may contribute to the auditory comprehension impairments of aphasics, they do not appear to be the primary cause of such impairments.

See also PHONOLOGICAL ANALYSIS OF LANGUAGE DISORDERS IN APHASIA.

—Sheila E. Blumstein

References


**Further Readings**


**Poverty: Effects on Language**

More than one in five children in the United States live in poverty, with pervasive consequences for their health and development. These consequences include effects on language. Children who live in poverty develop language at a slower pace than more advantaged children, and, after the age at which all children can be said to have acquired language, they differ from children from higher income backgrounds in their language skills and
manner of language use. Children from low-income families are also overrepresented among children diagnosed as language-impaired.

These relations between poverty and language are robust, but they are complicated to describe and difficult to interpret. Poverty itself is not a homogeneous condition but occurs on a gradient and is usually associated with other variables that affect language, particularly education and ethnicity. Thus, the source of poverty effects is sometimes obscure. Further, the language children display is always a combined function of their language skill and language style. Thus, the nature of poverty effects is sometimes unclear. This article describes the observed associations between poverty and language development, and then considers why those associations occur. Much of the relevant literature does not use poverty per se as a variable but instead uses correlated variables such as education or a composite measure of socioeconomic status (SES). Thus, an important task in studying poverty effects is to identify the causal factors at work when poverty, low levels of parental education, and other associated factors are correlated with low levels of language achievement in children.

**Language Differences Related to Income or SES.** Before speech begins, infants who live in poverty produce speech sounds and babble in much the same way and on the same developmental timetable as all normally developing infants (Oller et al., 1995), and even at 3 years of age, SES is unrelated to the accuracy of children’s articulation (Dollaghan et al., 1999). In contrast, virtually every other measure of language development reveals differences between children from low-income or low-SES families and children from higher income or higher SES families.

The clearest and largest SES-related difference is in the area of vocabulary. Recordings of spontaneous speech, maternal report measures, and standardized tests all show that children from low-income and low-SES families have smaller vocabularies than children from higher income and higher SES families (Rescorla, 1989; Hart and Risley, 1995; Dollaghan et al., 1999). By the age of 3 years, children from low-income families have productive vocabularies averaging around 500 words, while children from higher income families have productive vocabularies averaging more than 1000 words (Hart and Risley, 1995). Eighty percent of toddlers from low-income families score below the 50th percentile on the MacArthur Communicative Development Inventory (CDI) (Arriaga et al., 1998). Some findings suggest that low family income is more associated with measures of children’s productive vocabulary than with measures of their receptive vocabulary (Snow, 1999). Two studies have reported SES-related differences in children’s vocabulary, with lower SES children showing larger vocabularies. Both used the CDI, and both attributed their findings to lower SES mothers’ tendencies to overestimate their children’s abilities (Fenson et al., 1994; Feldman et al., 2000).

Grammatical development is also related to income or SES. Compared with higher SES children, children from lower social strata produce shorter responses to adult speech (McCarthy, 1930), score lower on standardized tests that include measures of grammatical development (Morrisset et al., 1990; Dollaghan et al., 1999), produce less complex utterances in spontaneous speech as toddlers (Arriaga et al., 1998) and at age 5 (Snow, 1999), and differ significantly on measures of productive and receptive syntax at age 6 (Huttenlocher et al., 2002). As an indicator of the magnitude of these effects on grammatical development, the low-income sample studied by Snow (1999) had an average MLU at age 3 years 9 months that would be typical of children more than a year younger, according to norms based on a middle-class sample. At age 5 years 6 months they had an average MLU typical of middle-class children age 3 years 1 month. On the other hand, the SES-related differences in productive syntax are not in whether children can or cannot use complex structures in their speech, but in the frequency with which they do so (Tough, 1982; Huttenlocher et al., in press). Studies of school-age children find SES-related differences in the communicative purposes to which language is put, such that children with less educated parents less frequently use language to analyze and reflect, to reason and justify, or to predict and consider alternative possibilities than children with more educated parents. The structural differences in children’s language associated with SES may be a by-product of these functional differences (Tough, 1982).

SES-related differences in school-age children also appear in the ability to communicate meaning through language and to draw meaning from language, sometimes referred to as speaker and listener skills (Lloyd, Mann, and Peers, 1998). In the referential communication task, which requires children to describe one item in an array of objects so that a visually separated listener with the same array can identify that item, lower SES children are less able than higher SES children to produce sufficiently informative messages and to use information in messages addressed to them to make correct choices (Lloyd et al., 1998). Children from lower socioeconomic strata also perform less well than higher SES children in solving mathematics word problems. This poorer performance reflects a difference in language ability, not mathematical ability, because the same children show no difference in performance in math calculations (Jordan, Huttenlocher, and Levine, 1992).

The foregoing effects are effects of poverty or SES on variation within the normal range. On average, children from low-income families acquire language at a slower rate and demonstrate both differences in language use and poorer language skills than children from higher income families. Low SES is also a correlate of the diagnosis of specific language impairment (SLI) (Tomblin et al., 1997), although it is not clear what this means, given that SLI is defined in terms of delay relative to norms and that normative development is slower for lower SES children (Fazio, Naremore, and Connell, 1996). For this reason, there have been calls for sensi-
tivity to SES effects on standardized measures in diagnosing SLI (Arriaga et al., 1998; Feldman et al., 2000), and efforts are underway to develop tests of language impairment that are direct tests of language learning ability rather than norm-referenced comparisons (Fazio, Naremore, and Connell, 1996; Campbell et al., 1997; Seymour, 2000).

Understanding the Relation Between Poverty and Language Development. Low family income cannot directly cause the depressed language skills associated with poverty but must operate via mediators that affect language. Potential mediators, or pathways, through which poverty operates may include factors with general effects on health and development, such as nutrition, exposure to environmental hazards, and quality of schools and child care, and may also include factors with specific effects on language, such as the opportunity for one-to-one contact with an adult (McCartney, 1984) and the language use of parents and classroom teachers (Hutenlocher et al., 2002).

A variety of evidence suggests that an important mediator of the relation between SES and language development in children is the nature of the language-learning environment provided by the family (Hoff, 2003) and that low levels of parental education, more than low income per se, affect the language-learning environment in the home. Two predictors of language development, the talk parents address to children and the exposure to books that parents provide, differ as a function of parental education. Compared with children of more educated parents, children of less educated parents hear less speech, hear less richly informative speech, receive less support for their own participation in conversation, and are read to less (Hart and Risley, 1995; Hoff-Ginsberg, 1998; U.S. Department of Education, 1998; Hoff, Laursen, and Tardif, 2002). When mediators of the relation between family SES and child language have been examined, properties of children’s language-learning environments have been found to account for most of the variance attributable to SES both for syntax and for lexical development (Hoff, 2003; Hutenlocher et al., 2002). In fact, even variation within low-income samples is attributable to variation in language-learning environments. Among a group of 5-year-olds in Head Start programs (thus, children from low-income families), standard scores on a measure of comprehension vocabulary (the Peabody Picture Vocabulary Test) were significantly related to maternal use of sophisticated vocabulary (i.e., low-frequency words) and to the frequency of supportive mother-child interactions (Weizman and Snow, 2001). In a different sample of 4-year-olds in Head Start, variation in productive and comprehension vocabulary was accounted for by children’s literacy experiences at home (Payne, Whitehurst, and Angell, 1994).

With respect to understanding the role of poverty in communication disorders, the literature presents a paradox. Poverty is associated with slow language development because poverty is associated with less supportive language-learning environments for children than more affluent situations. Poverty is also associated with the diagnosis of SLI, although most evidence suggests that the language environment is not the cause of SLI (Lederberg, 1980). In fact, studies of the heritability of language find a higher heritability for language impairment than for variation in language development within the normal range (Eley et al., 1999; Dale et al., 2000). If input is the mediator of poverty effects on language but input does not explain SLI, then why is poverty associated with SLI? The inescapable conclusion is that children can differ from the normative rate of development for two reasons: an impairment in the ability to learn language or an inadequate language-learning environment. Low percentile scores on measures of language development do not, by themselves, distinguish between these. While poverty does not cause language impairment or a communication disorder, the impoverished language-learning environment often associated with poverty can similarly impede language development.

See also SPECIFIC LANGUAGE IMPAIRMENT IN CHILDREN.

—Erika Hoff

References


Further Reading


Pragmatics

Pragmatics may be defined as “the study of the rules governing the use of language in social contexts” (McTear and Conti-Ramsden, 1992, p. 19). Although there is some debate as to what should be included under the heading of pragmatics, traditionally it has been thought to incorporate behaviors such as communicative intent (speech acts), conversational management (turn taking, topic manipulation, etc.), presuppositional knowledge, and culturally determined rules for linguistic politeness. Some authors, working from a framework in which pragmatics is seen as the motivating force behind other components of language such as syntax and semantics, include an expanded list of behaviors within this domain. For example, from this latter perspective, behaviors such as those occurring in the interactive context of early language acquisition would be considered pragmatic. Despite difficulty in determining where the boundaries of pragmatics should be drawn, as one considers how language is used in real interactions with other people, it is impossible not to cross over into areas more typically seen as social or cultural rather than linguistic (Ninio and Snow, 1996).

Although the social implications of impaired communication skills have been considered for some time, the pragmatic aspects of language impairment did not become a serious topic of study until the mid-1970s, following the lead of researchers studying typical language acquisition (see Leonard, 1998). The innovations for language assessment and intervention that stemmed from this work motivated Duchan (1984) to characterize these efforts as “the pragmatics revolution.” Gallagher (1991) summarized the contributions of pragmatics to assessment by noting that greater awareness of the pragmatic aspects of language resulted in a larger set of behaviors on which a diagnosis of language impairment could be made. It also highlighted the importance of contextual variables in spontaneous language production. Clinicians gained an understanding that controlling or standardizing these variables would fundamentally alter the nature of the interaction.

With respect to language intervention, goals were expanded to include a wider range of pragmatic behaviors. Further, providing intervention in more naturalistic contexts, thereby allowing communication to be motivated and reinforced by natural consequences, was highlighted. At the same time the value of using routines, scripts, and similar procedures to provide greater contextual support for language usage also gained favor among clinicians (Gallagher, 1991).

The study of pragmatics also brought insights into how communication might be linked to other aspects of
Problematic behaviors stemming from difficulty with symbol use include the often cited reliance on reenactment strategies (e.g., a preschool child who used the phrase “do ah” to mean that he was not feeling well, stemming from appropriate use in a prior context), as well as developing maladaptive ways of communicating to compensate for a lack of more conventional means (e.g., using head banging to communicate the desire to avoid an unpleasant task) (Wetherby, Prizant, and Schuler, 2000). Both of these examples of language use have important pragmatic implications.

Whereas it is accepted that pragmatic difficulties constitute a basic problem for children with ASD, the situation is less clear for children with SLI. It is well established that children in the latter group have difficulty with aspects of syntactic, morphological, and semantic development. Studies examining pragmatic behaviors have yielded more equivocal findings. For example, children with SLI have been found to be less capable of responding to stacked requests for clarification (Brinton, Fujiki, and Sonnenberg, 1988), less able to initiate utterances in conversation (Conti-Ramsden and Friel-Patti, 1983), and less adept at entering ongoing conversations (Craig and Washington, 1993) than typically developing peers at the same language level. Other researchers, however, have found that children with SLI performed similarly to typically developing peers on pragmatic variables when language level was controlled (e.g., Fey and Leonard, 1984; Leonard, 1986). In summarizing this work, it appears that many (but not all) children with SLI have difficulty with many (but not all) aspects of pragmatic language behavior. For some of these children, pragmatic deficits stem from problems with language form and content. For other children, pragmatic impairment is a central component of their language difficulty.

One way in which researchers have addressed the variability noted above has been to place children with SLI into subgroups more specifically characterizing the nature of their impairment. Several groups of researchers (e.g., Bishop and Rosenbloom, 1987; Conti-Ramsden, Crutchley, and Botting, 1997) have developed taxonomies identifying a group of children whose language impairments are pragmatic in nature. Labeled as “semantic pragmatic deficit syndrome,” these children are described as having a variety of pragmatic problems in the face of relatively typical structural language skills. Areas of deficit include inappropriate topic manipulation, difficulty assessing shared information, an overly high level of verbosity, and a lack of responsiveness to questions. Word-finding problems and difficulty comprehending language are also associated with this subgroup. In work by Conti-Ramsden, Crutchley, and Botting (1997), semantic pragmatic deficit syndrome characterized approximately 10% of 242 participants with SLI. In follow-up work, Conti-Ramsden and Botting (1999) found that although the subcategories of impairment were stable over the course of a year, many individual children moved between subcategories.

It might be argued that children with pragmatic problems in the face of relatively good structural language (and who do not meet criteria for ASD) might form a separate category of impairment. Bishop (2000) argued that this is not an ideal solution because pragmatic impairment is not limited to children with good structural skills, nor is it always found in association with semantic difficulties. Rather, it may be more productive to view structurally based SLI and ASD as two ends of a continuum on which many combinations of pragmatic and structural language deficits may occur.

In summary, pragmatic impairment may be found in a wide range of children with language problems. In
some cases, it may constitute a central component of the impairment. For other children, the problem may be secondary to other types of language problems. From a clinical standpoint, it is important to recognize that children with a variety of disabilities may have pragmatic problems. Given the close link between language and social behavior, it is perhaps as important to recognize that even language impairments that do not involve pragmatic behaviors are likely to have implications for social interaction. To be most productive, interventions should be structured not only to improve specific language skills but also to facilitate the use of language in interactions to improve relationships with peers and adults in the child’s social world.

—Martin Fujiki and Bonnie Brinton

References


Further Readings


Brinton, B., Fujiki, M., Winkler, E., and Loeb, D. (1986). Responses to requests for clarification in linguistically nor-
The onset of intentional communication late in the first year of life marks an infant’s active entry into his or her culture and ignites important changes in how others regard and respond to the infant. A significant delay in the onset of intentional communication is a strong indicator that the onset of productive language also will be delayed (McCaithren, Warren, and Yoder, 1996; Calandrella and Wilcox, 2000). Such a delay may hold the infant in a kind of developmental limbo because the onset of intentional communication triggers a series of transactional processes that support the emergence of productive language just a few months later. In this article we discuss the research on the effects of prelinguistic communication interventions aimed at teaching infants and toddlers with developmental delays to be clear, frequent prelinguistic communicators.

The onset of coordinated attention occupies a “pivotal” juncture in prelinguistic communication development. Before the emergence of coordinated attention, an infant’s intention is very difficult to discern (Bates, Benigni, Bretherton et al., 1979). Almost simultaneously with the emergence of coordinated attention, the child begins to move from preintentional to intentional communication. Requesting and commenting episodes provide the earliest contexts in which intentionality is demonstrated (Bates, O’Connell, and Shore, 1987). Both functions require the infant to shift his or her attention between his or her partner and an object. Requesting (also termed imperatives and protoimperatives in the literature) is commonly defined as behavior that clearly indicates that the child wants something. Commenting (also termed joint attention, indicating, declarative, and referencing in the literature) is the act of drawing another’s attention to or showing a positive affect about an object or interest (Bates, Benigni, Bretherton et al., 1987). Although other communicative functions also emerge during this period (e.g., greeting, protesting), requesting and commenting are considered the fundamental pragmatic building blocks of both prelinguistic and linguistic communication (Bruner, Roy, and Ratner, 1980). They are also the two most frequent functions expressed during the prelinguistic period (Wetherby, Cain, Yo-nclas et al., 1988).
The Transactional Model of Development and Intervention

The effectiveness of prelinguistic intervention in enhancing later language development depends on the operation of a transactional model of social communication development (Sameroff and Chandler, 1975; McLean and Snyder-McLean, 1978). The model presumes that early social and communication development is facilitated by bidirectional, reciprocal interactions between children and their environment. For example, a change in the child such as the onset of intentional communication may trigger a change in the social environment, such as increased linguistic mapping (i.e., naming objects and actions that the child is attending to) by their caregivers. These changes then support further development in the child (e.g., increased vocabulary), and subsequently further changes by the caregivers (e.g., more complex language interaction with the child). In this way, both the child and the environment change over time and affect each other in reciprocal fashion (e.g., more complex language interaction with the child).

A transactional model may be particularly well suited to understanding social-communication development in young children because caregiver-child interaction can play such an important role in this process. The period of early development (age birth to 3 years) may represent a unique time during which transactional effects can have a substantial impact on development. Young children’s relatively restricted repertoire during this early period can make changes in their behavior more salient and easily observable to caregivers. This in turn may allow adults to be more specifically contingent with their responses to developing skills of the child than is possible later in development, when children’s behavioral repertoire is far more expansive and complex. During this natural window of opportunity, the transactional model may be employed by a clever practitioner to multiply the effects of relatively circumscribed interventions and perhaps alter the very course of the child’s development in a significant way. But the actions of the practitioner may need to be swift and intense, or they may be muted by the child’s steadily accumulating history.

The generation of strong transactional effects in which the growth of emotional, social, and communication skills is scaffolded by caregivers can have a multiplier effect in which a relatively small dose of early intervention may lead to long-term effects. These effects are necessary when we consider that a relatively “intensive” early intervention by a skilled clinician may represent only a few hours a week of a young child’s potential learning time (e.g., 5 hours per week of intensive interaction would represent just 5 percent of the child’s available social and communication skill learning time if we assume the child is awake and learning 100 hours per week). Thus, unless direct intervention accounts for a large portion of a child’s waking hours, transactional effects are mandatory for early intervention efforts to achieve their potential.

Effects of Prelinguistic Communication Intervention

In their initial explorations of the effects of prelinguistic communication intervention, Yoder and Warren demonstrated that increases in the frequency and clarity of prelinguistic requesting by children with developmental delays as a result of the intervention covaried with substantial increases in linguistic mapping by teachers and parents who were naive as to the specific techniques and goals of the intervention (Warren, Yoder, Gazdag et al., 1993; Yoder, Warren, Kim et al., 1994). These studies also demonstrated strong generalization effects in that the intentional requesting function they taught was shown to generalize across people, setting, communication styles, and time. Based on the promising results of these initial studies, Yoder and Warren (1998, 1999a, 1999b, 2001a, 2001b) conducted a relatively large (N = 58) longitudinal experimental study of the effects of prelinguistic communication intervention on the communication and language development of children with general delays in development. This study represented an experimental analysis of the transactional model of early social communication development.

Fifty-eight children between the ages of 17 and 32 months (mean, 23; SD, 4) with developmental delays and their primary parent participated in this study. Fifty-two of these children had no productive words at the outset of the study; the remaining six children had between one and five productive words. All children scored below the 10th percentile on the expressive scale of the Communication Development Inventory (Fenson et al., 1991).

The children were randomly assigned to one of two treatment groups. Twenty-eight of the children received an intervention termed “prelinguistic milieu teaching” (PMT) for 20 minutes per day, 3 or 4 days per week, for 6 months. The other 30 children received an intervention termed “responsive small group” (RSG). PMT represented an adaptation of milieu language teaching (e.g., Warren and Bambara, 1989; Warren, 1992). RSG represented an adaptation of the responsive interaction approach (Wilcox, 1992; Wilcox and Shannon, 1998). These interventions are described in detail elsewhere (e.g., Warren and Yoder, 1998; Yoder and Warren, 1998). Caretakers were kept naive as to the specific methods, measures, records of child progress, and child goals throughout the study. This allowed Yoder and Warren to investigate how change in the children’s behavior as a result of the interventions might affect the behavior of the primary caretaker, and how this in turn might affect the child’s development later in time. Data were collected at five points in time for each dyad: at pretreatment, at post-treatment, and 6, 12, and 18 months after the completion of the intervention.

Both interventions had generalized effects on intentional communication development. However, the treatment that was most effective depended on the pre-treatment maternal interaction style and the education level of the mother (Yoder and Warren, 1998, 2001b).
Specifically, Yoder and Warren found that for children of highly responsive, relatively well-educated mothers, PMT was effective in fostering intentional communication development. However, for children with relatively unresponsive mothers, RSG was relatively more successful in fostering generalized intentional communication development.

The two interventions differ along a few important dimensions that provide a plausible explanation for these effects. PMT uses a child-centered play context in which communication prompts for more advanced forms of communication are employed, as well as social consequences for target responses such as specific acknowledgment and compliance. RSG emphasizes following the child’s attentional lead and being highly responsive to child initiation while avoiding the use of direct prompts for communication. Maternal interaction style may have influenced which intervention was most beneficial, because children may develop expectations concerning interactions with adults (including teachers and interventionists) based on their history of interaction with their primary caretaker(s). Thus, children with responsive parents may learn to persist in the face of communication breakdowns, such as might be occasioned by a direct prompt or time delay, because their history leads them to believe that their communication attempts will usually be successful. On the other hand, children without this history may cease communicating when their initial attempt fails. Thus, children of responsive mothers in the PMT group persisted when prompted and thus learned effectively in this context, while children with unresponsive parents did not. But when provided with a highly responsive adult who virtually never prompted them over a 6-month period, children of unresponsive mothers showed greater gains than children of responsive parents receiving the same treatment.

The effects of maternal responsivity as a moderator of intervention effects rippled throughout the longitudinal follow-up period. Yoder and Warren demonstrated that children in the PMT group with relatively responsive mothers received increased amounts of responsive input in direct response to their increased intentional communication (Yoder and Warren, 2001b). Furthermore, the effects of the intervention were found on both protoimperatives and protodeclaratives (Yoder and Warren, 1999a), became greater with time, and impacted expressive and receptive language development 6 and 12 months after intervention ceased (Yoder and Warren, 1999b, 2001a). It is important to consider this finding in light of the substantial number of early intervention studies in which the effects were reported to wash out over time (Farren, 2000). Finally, the finding that amount of responsive input by the primary caregiver was partly responsible for the association between intentional communication increases and later language development (Yoder and Warren, 1999a), coupled with the longitudinal relationship between maternal responsivity and expressive language development (Yoder and Warren, 2001a), supports the prediction of the transactional model that children’s early intentional communication will elicit mother’s linguistic mapping, which in turn will facilitate the child’s vocabulary development.

Conclusion

Prelinguistic communication intervention represents a promising approach for young children with developmental delays. Research on this approach has been quite limited to date. However, it is clear that the effectiveness of specific interventions may be dependent to some extent on mediating effects of caretaker responsivity. Therefore, the combining specific child-centered techniques such as PMT with parent training aimed at enhancing caretaker responsivity may be the most efficacious approach.

See also COMMUNICATION DISORDERS IN INFANTS AND TODDLERS.

—Steven F. Warren and Paul J. Yoder

References


Preschool Language Intervention

In 2000, the British Medical Journal published an important clinical trial of preschool speech and language services in 16 community clinics in Bristol, England (Glogowska et al., 2000). This study was the largest trial of its kind to address the communication problems of preschool children. Unfortunately, its findings were largely negative; after 12 months, the treatment group showed a significant advantage on only one of the five primary outcome variables, auditory comprehension, over a control group that had received only “watchful waiting” over the same time period. This lone treatment effect was small and, arguably, not clinically significant. Most children in both groups were still eligible for preschool speech and language services at the end of the 12-month study period.

As disappointing as these results are, it is important to note that the study was designed to study the effectiveness of early communication services as typically provided in one community in the United Kingdom (Law and Conti-Ramsden, 2000). Effectiveness studies evaluate treatment effects under relatively typical clinical conditions. As such, the investigators learned that, on average, the children in the treatment group received only 6.2 hours of intervention, or 30 minutes per month. In fact, the most intervention provided to any child was only 15 hours over a 12-month period! This study demonstrates most clearly that small, insignificant doses of early language intervention are not effective in eliminating or reducing the broad range of problems associated with preschool language impairment (see Specific Language Impairment in Children; Social Development and Language Impairment: Preschool Language Intervention).

Unlike studies of effectiveness, which monitor client change under typical clinical conditions, efficacy studies are designed to determine, under more idealized, laboratory conditions, whether an intervention is directly responsible for positive outcomes. There is ample evidence that when preschool language interventions are applied regularly with reasonable intensity, they are efficacious, leading to clinically significant improvement in the children’s language and early literacy skills. There are many varieties of preschool interventions, however, and clinicians must carefully consider the options available to them.

The principal differences between different preschool intervention approaches are best captured by determining where the interventions fall on a continuum of intrusiveness. Approaches that are highly intrusive use direct teaching methods in clinical settings, usually with the clinician as the intervention agent, to address predetermined treatment objectives, such as specific words or grammatical structures. In contrast to the prescriptive character of highly intrusive approaches, minimally intrusive approaches have goals that are stated more broadly, with less focus on specific targets. Example targets include the use of longer and more complex sentences, personal reports, and stories, with the child using an increasingly varied vocabulary. This gives the child latitude to learn from the rich set of linguistic options available in intervention contexts. In general, the clinician exerts limited control over the child’s agenda. Descriptions and examples of approaches at each end of the intrusiveness continuum follow.

In protocols that are maximally intrusive, the child may examine a picture, object, or event presented by a clinician, who then presents a linguistic model and a request for the child to imitate. If the child imitates correctly, the clinician provides social or other reinforcement and then presents another stimulus set. If the child imitates incorrectly, the stimulus is repeated or sim-
child to complete (e.g., “Tell Sandy you want a cookie”) or by altering its form in some way (“Can I have a cookie now?”). In interventions such as this, it is easy and appropriate to focus on early literacy skills, such as letter knowledge, rhyming, and phonological awareness. In keeping with the limited clinician intrusiveness, however, the clinician does not directly teach specific words, language structures, story structure, or early literacy targets, nor are efforts made to get the child to imitate or to produce language out of context.

As appealing as these child-oriented approaches are, there is only limited empirical evidence that they are efficacious in facilitating language use among children with language impairments. Furthermore, it has not been adequately demonstrated that focusing broadly on the communication of meaning leads to gains in the specific areas of grammatical, phonological, and discourse weakness exhibited by preschoolers with language impairment. Techniques such as following the child’s lead, recasting the child’s utterances, and following the child’s utterances with open-ended questions can be efficiently taught to parents or paraprofessionals, and this is an important feature. For example, Dale et al. (1996) taught parents to use these procedures during shared book reading with their children over two relatively brief sessions. Parents made more changes as a result of the intervention than the children did, but outcomes were measured after only 2 months. A longer intervention period may have resulted in greater effects on the child’s performance.

Contemporary language interventions typically are hybrids that fall somewhere between the extremes in intrusiveness. For example, so-called milieu interventions blend the identification of discrete intervention targets and direct teaching using imitation and other prompts (i.e., more intrusive components) with the principles of creating natural contexts for communication, following the child’s lead, and recasting the child’s utterances (i.e., less intrusive components). These approaches appear to be especially efficacious for children at the single-word or early multiword stages (Yoder, Kaiser, and Alpert, 1991). Gibbard (1994) demonstrated that parents can be taught to use milieu procedures in as few as 11 sessions over a 6-month period, yielding effects commensurate with those of clinician-administered treatment. When they are applied with moderate intensity, milieu approaches increase not only the length and complexity of children’s utterances, but also the children’s conversational assertiveness and responsiveness (Warren, McQuarter, and Rogers-Warren, 1984).

Another popular hybrid intervention is called focused stimulation. Most focused stimulation approaches create contexts within which the interventionist produces frequent models of the child’s social and linguistic targets and creates numerous opportunities for the child to produce them. Interventionists follow the child’s lead, recasting the child’s utterances and using the child’s language targets, but they do not prompt the child to imitate. Fey et al. (1993) used this type of approach over a 5-month period to facilitate the grammatical abilities of a group of 4- to 6-year-old children with impairments of grammatical production. They also trained parents to use these techniques over a 12-session parent intervention. The children who received intervention exclusively
from their parents made gains that were, on average, equivalent to the gains of the children who received 3 hours of weekly individual and group intervention from a clinician. Observed gains in the parent group were not as consistent across children as were the gains of the children in the clinician group, however.

In sum, preschool language intervention of several different types can be efficacious. Although individual clinicians have their strong personal preferences, there are few experimental indications that any one approach is dramatically superior to the others. To achieve clinically meaningful effects, however, these interventions must be presented rigorously over periods of at least several months. Furthermore, it remains unclear whether existing approaches are sufficient to minimize the risks for later social, behavioral, and academic problems preschoolers with language impairments typically experience once they reach school. To this end, promising hybrid preschool classroom interventions have been developed that aim to enhance not only the children’s spoken language, but their problems in social adaptation and early literacy as well (Rice and Wilcox, 1995; van Kleeck, Gillam, and McFadden, 1998).

—Marc E. Fey

References


Further Readings


Among the sequelae of certain types of brain damage are impairments in the production and perception of speech prosody. Prosody serves numerous functions in language, including signaling lexical differences when used phonemically in tone languages, providing cues to stress, sentence type or modality, and syntactic boundaries, and conveying a speaker’s emotions. Any or all of these functions of prosody may be impaired subsequent to brain damage. The hemispheric lateralization of the brain lesion seems to play an important role in the nature of the ensuing deficits; however, the neural substrates for prosody are still far from clear (see Baum and Pell, 1999, for a review). Historically, clinical impressions led to the contention that subsequent to right hemisphere damage (RHD), patients would present with flat affect and monotonous speech, whereas patients with left hemisphere damage (LHD) would maintain normal speech prosody. As research progressed, several alternative theories concerning the control of speech prosody were posited. Among these are the hypothesis that affective or emotional prosody is controlled within the right hemisphere, and thus RHD would yield emotional prosodic deficits, whereas linguistic prosody is controlled within the left hemisphere, yielding linguistic prosodic deficits when damage is confined to the left hemisphere (e.g., Van Lancker, 1980; Ross, 1981). A second hypothesis proposes that prosody is principally controlled in subcortical regions and via cortical-subcortical connections (e.g., Cancelliere and Kertesz, 1990); evidence of prosodic deficits in individuals with Parkinson’s disease supports this view. A third alternative contends that the individual acoustic cues to prosody (i.e., duration, amplitude, and fundamental frequency) are differentially lateralized to the right and left hemispheres, with temporal properties processed by the left hemisphere and spectral properties by the right hemisphere (e.g., Van Lancker and Sidis, 1992). Whereas several recent investigations have utilized functional neuroimaging techniques in normal individuals to address these hypotheses (e.g., Gandour et al., 2000), by far the most data have been gathered in studies of individuals who have suffered brain damage. These investigations allow us to characterize the nature of prosodic deficits that may emerge in neurologically impaired populations. The discussion is divided into affective and linguistic prosodic impairments.

Beginning with deficits in the production and perception of affective prosody, one of the salient speech characteristics of individuals who have suffered RHD is a flat affect. That is, in conjunction with a reduction in emotional expression as reflected in facial expressions, clinical impressions suggest that individuals with RHD tend to produce speech that is reduced, if not devoid, of affect. In fact, based on clinical judgments of the speech of RHD patients with varying sites of lesion, Ross (1981) proposed a classification system for affective impairments, or aprosodias, that paralleled the popular aphasia syndrome classification system of Goodglass and Kaplan (1983). Ross’s 1981 classification scheme sparked a good deal of research on affective prosodic deficits that ultimately resulted in its abandonment by the majority of investigators. However, the investigations it catalyzed contributed significantly to our understanding of prosodic impairments; much of the work inspired by Ross’s proposal took advantage of increasingly reliable methods such as acoustic analysis of speech. When studying RHD patients in an acute stage, results seemed to support impairments in patients’ ability to accurately signal emotions such as happiness, sadness, and anger.

**Prosodic Deficits**


Robertson, S. B., and Ellis Weismer, S. (1997). The influence of sentence type or modality, and syntactic boundaries, on the production and perception of speech prosody. *Prosody serves numerous functions in language, including signaling lexical differences when used phonemically in tone languages, providing cues to stress, sentence type or modality, and syntactic boundaries, and conveying a speaker’s emotions. Any or all of these functions of prosody may be impaired subsequent to brain damage. The hemispheric lateralization of the brain lesion seems to play an important role in the nature of the ensuing deficits; however, the neural substrates for prosody are still far from clear (see Baum and Pell, 1999, for a review). Historically, clinical impressions led to the contention that subsequent to right hemisphere damage (RHD), patients would present with flat affect and monotonous speech, whereas patients with left hemisphere damage (LHD) would maintain normal speech prosody. As research progressed, several alternative theories concerning the control of speech prosody were posited. Among these are the hypothesis that affective or emotional prosody is controlled within the right hemisphere, and thus RHD would yield emotional prosodic deficits, whereas linguistic prosody is controlled within the left hemisphere, yielding linguistic prosodic deficits when damage is confined to the left hemisphere (e.g., Van Lancker, 1980; Ross, 1981). A second hypothesis proposes that prosody is principally controlled in subcortical regions and via cortical-subcortical connections (e.g., Cancelliere and Kertesz, 1990); evidence of prosodic deficits in individuals with Parkinson’s disease supports this view. A third alternative contends that the individual acoustic cues to prosody (i.e., duration, amplitude, and fundamental frequency) are differentially lateralized to the right and left hemispheres, with temporal properties processed by the left hemisphere and spectral properties by the right hemisphere (e.g., Van Lancker and Sidis, 1992). Whereas several recent investigations have utilized functional neuroimaging techniques in normal individuals to address these hypotheses (e.g., Gandour et al., 2000), by far the most data have been gathered in studies of individuals who have suffered brain damage. These investigations allow us to characterize the nature of prosodic deficits that may emerge in neurologically impaired populations. The discussion is divided into affective and linguistic prosodic impairments.

Beginning with deficits in the production and perception of affective prosody, one of the salient speech characteristics of individuals who have suffered RHD is a flat affect. That is, in conjunction with a reduction in emotional expression as reflected in facial expressions, clinical impressions suggest that individuals with RHD tend to produce speech that is reduced, if not devoid, of affect. In fact, based on clinical judgments of the speech of RHD patients with varying sites of lesion, Ross (1981) proposed a classification system for affective impairments, or aprosodias, that paralleled the popular aphasia syndrome classification system of Goodglass and Kaplan (1983). Ross’s 1981 classification scheme sparked a good deal of research on affective prosodic deficits that ultimately resulted in its abandonment by the majority of investigators. However, the investigations it catalyzed contributed significantly to our understanding of prosodic impairments; much of the work inspired by Ross’s proposal took advantage of increasingly reliable methods such as acoustic analysis of speech. When studying RHD patients in an acute stage, results seemed to support impairments in patients’ ability to accurately signal emotions such as happiness, sadness, and anger.

**Prosodic Deficits**


Robertson, S. B., and Ellis Weismer, S. (1997). The influence of sentence type or modality, and syntactic boundaries, on the production and perception of speech prosody. *Prosody serves numerous functions in language, including signaling lexical differences when used phonemically in tone languages, providing cues to stress, sentence type or modality, and syntactic boundaries, and conveying a speaker’s emotions. Any or all of these functions of prosody may be impaired subsequent to brain damage. The hemispheric lateralization of the brain lesion seems to play an important role in the nature of the ensuing deficits; however, the neural substrates for prosody are still far from clear (see Baum and Pell, 1999, for a review). Historically, clinical impressions led to the contention that subsequent to right hemisphere damage (RHD), patients would present with flat affect and monotonous speech, whereas patients with left hemisphere damage (LHD) would maintain normal speech prosody. As research progressed, several alternative theories concerning the control of speech prosody were posited. Among these are the hypothesis that affective or emotional prosody is controlled within the right hemisphere, and thus RHD would yield emotional prosodic deficits, whereas linguistic prosody is controlled within the left hemisphere, yielding linguistic prosodic deficits when damage is confined to the left hemisphere (e.g., Van Lancker, 1980; Ross, 1981). A second hypothesis proposes that prosody is principally controlled in subcortical regions and via cortical-subcortical connections (e.g., Cancelliere and Kertesz, 1990); evidence of prosodic deficits in individuals with Parkinson’s disease supports this view. A third alternative contends that the individual acoustic cues to prosody (i.e., duration, amplitude, and fundamental frequency) are differentially lateralized to the right and left hemispheres, with temporal properties processed by the left hemisphere and spectral properties by the right hemisphere (e.g., Van Lancker and Sidis, 1992). Whereas several recent investigations have utilized functional neuroimaging techniques in normal individuals to address these hypotheses (e.g., Gandour et al., 2000), by far the most data have been gathered in studies of individuals who have suffered brain damage. These investigations allow us to characterize the nature of prosodic deficits that may emerge in neurologically impaired populations. The discussion is divided into affective and linguistic prosodic impairments.

Beginning with deficits in the production and perception of affective prosody, one of the salient speech characteristics of individuals who have suffered RHD is a flat affect. That is, in conjunction with a reduction in emotional expression as reflected in facial expressions, clinical impressions suggest that individuals with RHD tend to produce speech that is reduced, if not devoid, of affect. In fact, based on clinical judgments of the speech of RHD patients with varying sites of lesion, Ross (1981) proposed a classification system for affective impairments, or aprosodias, that paralleled the popular aphasia syndrome classification system of Goodglass and Kaplan (1983). Ross’s 1981 classification scheme sparked a good deal of research on affective prosodic deficits that ultimately resulted in its abandonment by the majority of investigators. However, the investigations it catalyzed contributed significantly to our understanding of prosodic impairments; much of the work inspired by Ross’s proposal took advantage of increasingly reliable methods such as acoustic analysis of speech. When studying RHD patients in an acute stage, results seemed to support impairments in patients’ ability to accurately signal emotions such as happiness, sadness, and anger.
However, the majority of investigations of patients who had reached a more chronic stage (i.e., at least 3 months post onset) reported few differences between RHD patients and normal controls in signaling various emotions, as reflected in acoustic measures as well as perceptual judgments. Occasionally, studies have reported affective prosodic impairments in speech production subsequent to LHD (e.g., Cancelliere and Kertesz, 1990), although such findings are far less frequent.

With respect to the perception of affective prosody, early studies again suggested deficits in the processing of emotions cued by vocal signals subsequent to RHD (see Baum and Pell, 1999). Additional investigations have also indicated that LHD patients may exhibit deficits in the perception of affective prosody, particularly when the processing load is heavy (e.g., Tompkins and Flowers, 1985). The finding that both LHD and RHD may yield impairments in prosodic processing led to the proposal that the individual acoustic properties that serve as prosodic cues (i.e., duration, F0, and amplitude) may be processed independently in the two cerebral hemispheres and that patients with RHD and LHD may rely to different degrees on multiple cues (e.g., Van Lancker and Sidtis, 1992).

With regard to linguistic prosody, in keeping with the hypothesized functional lateralization of prosody described earlier (Van Lancker, 1980), numerous investigations have demonstrated that individuals with LHD exhibit impairments in the production of linguistic prosody, particularly at the phonemic level (i.e., in tone languages such as Mandarin, Norwegian, or Thai; e.g., Gandour et al., 1992). Deficits in the ability to signal emphatic stress contrasts, declarative versus interrogative sentence types, and syntactic clause boundaries have also been shown subsequent to LHD (e.g., Danly and Shapiro, 1982), but some studies have shown similar impairments in RHD patients (e.g., Pell, 1999) or have demonstrated that the production of only certain acoustic cues, primarily temporal parameters, is affected in LHD patients (e.g., Baum et al., 1997). The clearest evidence for the role of the left hemisphere in the production of linguistic prosody comes from studies of the phonemic use of tone; while this is arguably the “most linguistic” of the functions of prosody, it is also the smallest unit (i.e., a single syllable) in which prosodic cues may be manifest. It has therefore been suggested that the size or domain of the production unit may play a role in the brain regions implicated in prosodic processing. As an obvious corollary, patients with LHD and RHD may display impairments limited to different domains of prosodic processing.

Impairments in the perception or comprehension of linguistic prosody have also been found in both LHD and RHD patient groups, with varying results depending on the nature of the stimuli or the task. Investigations focusing on the perception of stress cues have mainly reported reduced performance relative to normal by individuals with LHD. For instance, several studies have shown that LHD patients are impaired in the ability to identify phonemic (lexical) and emphatic stress (e.g., Emmorey, 1987). With regard to the perception of linguistic prosodic cues at the phrase or sentence level, both individuals with RHD and LHD have difficulty identifying declarative, interrogative, and imperative sentence types on the basis of prosodic cues alone. LHD but not RHD patients tend to be relatively more impaired in linguistic than affective prosodic perception when direct comparisons are made within a single study (Heilman et al., 1984). Baum and colleagues (1997) have also noted impairments in the perception of phrase boundaries by both LHD and RHD patients. Investigations of individuals with basal ganglia disease due to Parkinson’s or Huntington’s disease have also reported deficits in the comprehension of prosody (e.g., Blonder, Gur, and Gur, 1989), suggesting that subcortical structures or cortical-subcortical connections are important in prosodic processing. Due to its multiple functions in language, understanding prosodic deficits and the neural substrates implicated in the processing of prosody is clearly a complex task.

Although this article has considered affective and linguistic prosody separately, this represents a somewhat artificial distinction, as they are integrated in natural speech production and perception. A handful of recent investigations have begun to address this integration, with mixed results appearing even in normal individuals (e.g., Pell, 1999). Exploring the integration of affective and linguistic prosody in individuals who have suffered brain damage only compounds the problem and the inconsistencies.

In summary, impairments in the production and perception of speech prosody may emerge subsequent to focal brain damage to numerous cortical and subcortical regions. The precise nature of the deficit may depend in part on the site of the lesion, but it seems to vary along the dimensions of the prosodic functional load (from affective to linguistic), the size or domain of the production or processing unit, and the specific acoustic parameters contributing to the prosodic signal. Prosodic deficits clearly interact with other communicative impairments, including disorders of linguistic and pragmatic processing, contributing to the symptom complexes associated with the aphasias, motor speech disorders, and right hemisphere communication deficits.

See also right hemisphere language and communicative function in adults; right hemisphere language disorders.

—Shari R. Baum

References


**Reversibility/Mapping Disorders**

Impaired comprehension of reversible sentences is widely observed in aphasia. Reversible sentences (e.g., *The cat chased the dog*) cannot be interpreted accurately without attention to word order and other syntactic devices, whereas the sole plausible interpretation of nonreversible sentences (e.g., *The cat drank the milk*) can be derived from content words via semantic or pragmatic inferencing. Impaired comprehension of reversible sentences, along with relatively intact comprehension of single words and nonreversible sentences, most frequently co-occurs with Broca’s aphasia but is also observed in other forms of aphasia (Caramazza and Zurif, 1976; Martin and Blossom-Stuch, 1986; Caramazza and Micelli, 1991). This comprehension pattern, termed asyntactic comprehension, has been studied intensively as evidence about the syntactic abilities of aphasic listeners. Accounts of asyntactic comprehension differ in two dimensions. The first is competence versus performance: Does the failure to interpret reversible sentences correctly derive from a loss of linguistic knowledge or language-processing ability, or does this failure stem from performance factors such as resource limitations? The second dimension is parsing versus mapping: Does asyntactic comprehension derive from a failure to parse or from a failure to map an accurate parse onto a semantic representation?

**Competence Versus Performance**

Early competence-based interpretations of asyntactic comprehension pointed to loss of linguistic knowledge or damage to the human parser as a common underlying source for this comprehension impairment and for *agrammatism*, a speech production pattern found in some Broca’s aphasics. *Agrammatism* is characterized by...
omission or misselection of grammatical morphemes and/or simplified, fragmentary grammatical structure. This hypothesis of a central syntactic disorder (Caramazza and Zurif, 1976; Berndt and Caramazza, 1980) was motivated by the observation that syntactic comprehension and agrammatism of speech both suggest a limited exploitation of syntactic devices. However, double dissociations between syntactic comprehension and agrammatic production (Goodglass and Menn, 1985) argue against a central account. Competence-based explanations—using the term “competence” broadly to apply to the absolute inability to perform specific linguistic operations as a result of either loss of knowledge or damage to the psychological mechanisms responsible for computing linguistic representations—are undermined by the findings that (1) syntactic comprehension can be induced in normal subjects under resource-demanding conditions (Miyake, Carpenter, and Just, 1994; Blackwell and Bates, 1995); (2) aphasic performance varies greatly from session to session (Kolk and van Grunsven, 1985) and from task to task (Cupples and Inglis, 1993); and (3) syntactic comprehenders frequently perform close to normally on grammaticality judgment tasks (Linebarger, Schwartz, and Saffran, 1983), detecting grammatical ill-formedness in the same structures that they do not reliably comprehend. These findings are more compatible with a performance account than with an account that implicates loss of the ability to perform the relevant language-processing operations even under optimal conditions.

Performance accounts differ regarding the nature of the hypothesized resource limitation. Some point to a global resource deficit (Blackwell and Bates, 1995); others invoke more specific limitations (Miyake et al., 1994; Caplan and Waters, 1995). Performance accounts also differ regarding the linguistic operations disrupted by the hypothesized resource limitation: some implicate parsing (Kolk, 1995); others point to subsequent interpretative processes.

Asyntactic Comprehension as Parsing Failure

This class of performance-based explanations posits a failure to retrieve the syntactic structure of the input sentence, a failure that may occur only when task demands are high (Frazier and Friederici, 1991; but see Linebarger, 1995). On the basis of patterns of performance on comprehension tasks, some investigators have attempted to pinpoint the grammatical locus of this failure, implicating, for example, the processing of closed-class elements (Bradley, Garrett, and Zurif, 1980), syntactically moved elements (Grodzinsky, 1990), or referential dependencies (Mauner, Fromkin, and Cornell, 1993). One difficulty for these more fine-grained hypotheses (see also Linebarger, 1995) is the variability observed in aphasric performance across different syntactic structures (Berndt, Mitchum, and Haendiges, 1996). For example, these accounts predict good performance on simple active sentences, and hence fail to account for the difficulties posed for some patients by such sentences (Schwartz, Saffran, and Marin, 1980). And even those aphasric patients who normally perform well on simple active reversible sentences may fail when the lexical content is manipulated so that the syntactically correct interpretation is the opposite of the interpretation supported by lexicosemantic heuristics, as in, for example, a semantic/pragmatic anomaly task requiring the subject to detect the anomaly of The cheese ate the mouse (Saffran, Schwartz, and Linebarger, 1998).

An additional point to note is that while the linguistically fine-grained accounts appeal to heuristics to explain observed patterns of comprehension, they invoke these heuristics only as a response to parsing failure. The Saffran et al. (1998) data, in contrast, suggest that heuristics may occur in parallel with, and sometimes in competition with, syntactic analysis. Such a view accords with studies of sentence processing in normals, where the influence of extragrammatical heuristics in normal sentence comprehension is well documented (Slobin, 1966; Bever, 1970; Trueswell, Tanenhaus, and Garney, 1994).

Asyntactic Comprehension as Mapping Failure

This explanation for asyntactic comprehension implicates the mapping between syntactic structure and semantic interpretation (Linebarger, Schwartz, and Saffran, 1983). On this account, asyntactic listeners do construct an adequate representation of the structure of the input sentence but fail to exploit this syntactic information for the recovery of meaning, specifically of thematic roles such as agent and theme. The frequently observed difficulty posed by passive and object-gapped sentences follows, on this account, from the fact that the order of content words in these structures conflicts with extragrammatical order-based heuristics (Caplan, Baker, and Dehaut, 1985). If such heuristics occur in parallel with grammatically based processing, as suggested by the literature on normals, then conflict between grammatical structure and extragrammatical heuristics would be predicted to lead to more errorful performance.

Evidence to choose between these two hypotheses is sparse. The grammaticality judgment data do not undermine the performance (as opposed to the competence) version of the parsing hypothesis, on the assumption that the grammaticality judgment task is less resource-demanding than the sentence-picture matching task or other paradigms that require not only syntactic analysis but also semantic interpretation. Parsing, on this account, is performed in optimal circumstances, but not when other task demands are high.

The grammaticality judgment data do not contradict the mapping account, either. This explanation for asyntactic comprehension posits a normal syntactic parse that is not adequately mapped onto an interpretation. In fact, patterns of performance observed within the grammaticality judgment task may be seen as supporting the mapping hypothesis. Errors related to constituent structure, verb subcategorization, and the legitimacy of syntactic gaps were detected more reliably by agrammatic patients than errors involving coindexation of pronouns and other referential elements (Linebarger, 1990); such patterns suggest an initial “first-pass” recovery of con-
stinent structure that is not fully interpreted, a pattern which falls naturally from the mapping hypothesis.

The mapping hypothesis also receives support from a study in which aphasic subjects judged the plausibility of simple reversible sentences and, in addition, of “padded” versions of these sentences (Schwartz et al., 1987; see also Kolk and Weijts, 1995). In the padded versions, the basic SVO (subject-verb-object) structure was elaborated with extraneous material. For example, subjects were presented with *The bird swallowed the worm*, its padded counterpart, *As the sun rose, the bird in the cool wet grass swallowed the worm quickly and went away*, and with the role-reversed versions of these sentences. In addition, these same predicate argument structures were embedded in noncanonical structures such as passives, object gaps, and other deviations from the simple SVO structure that typically cause difficulties for syntactic comprehenders. The agrammatic and conduction aphasics in this study performed well above chance on both the simple and padded sentences; their performance declined to chance or near chance only on the noncanonical structures. The good performance on padded sentences supports the view that syntactic comprehenders are able to construct an adequate representation of constituent structure, because the extraction of the elements critical to the plausibility judgment (*bird, swallowed, and worm*) requires an analysis of the structure of the sentence. A nonsyntactic “nearest NP” strategy, for example, would lead subjects to reject the padded sentence above, since *grass* immediately precedes *swallowed*.

The literature on mapping therapy, an approach to remediation based on this hypothesis, contains reports (Jones, 1986; Byng, 1988) of striking gains resulting from a training protocol focusing on the relationship between grammatical functions such as subject/object and thematic roles such as agent/theme. While subsequent studies have reported a variety of outcomes for this approach to therapy, the reported successes suggest that for at least a subset of patients, the breakdown in processing may occur in the assignment of thematic roles on the basis of grammatical function.

**Asyntactic Comprehension as Evidence About Language Processing**

It can be argued that the fragility of linguistic processing in aphasia, whatever its cause, results in a disproportionate influence of extralinguistic processing based on lexical content and word order rather than grammatical structure. Therefore the patterns of misinterpretation observed in aphasic subjects may not directly reflect their linguistic impairments, but rather a complex interaction between inefficient or inaccurate linguistic analysis and extragrammatical interpretative processes. Furthermore, the heterogeneous patterns of interpretive errors even within specific subgroups such as agrammatics suggest that there may be no unitary explanation for the impaired comprehension of reversible sentences in aphasia.

*See also ATTENTION AND LANGUAGE; MEMORY AND PROCESSING CAPACITY; TRACE DELETION HYPOTHESIS.*

—Marcia C. Linebarger

**References**


Right Hemisphere Language and Communication Functions in Adults

The role of the right cerebral hemisphere in language and communication represents a relatively young area of research that has grown rapidly since the late 1970s. Recent interest in the role of the right hemisphere reflects an emphasis on language as a tool for communication in natural contexts, and an awareness that normal language use is the product of many regions of the two hemispheres working in concert. Left hemisphere structures are routinely linked to the nuts and bolts of what might be termed basic language: phonology, lexical semantics, and syntax. In contrast, right hemisphere structures have been implicated in less tightly constrained domains, including some uses of prosody (the “melody” of speech), metaphor, discourse such as conversations, stories, indirect requests, and other forms of nonliteral language, and even the social-cognitive basis for discourse. These domains most closely associated with the right hemisphere are especially sensitive to context and are ideally suited to expressing nuance. This article will first present some general issues pertaining to research in these areas and then, in turn, representative findings relating to prosody, lexical processing and metaphor, and discourse.

Claims related to right hemisphere contributions to language and communication can be stated in a strong form: that a specific function is housed in some region of the right hemisphere that is necessary and sufficient to support that function. However, most claims are more general and also weaker: that normal task performance draws on intact right as well as left hemisphere structures. For example, understanding the point of an ironic comment rests on a listener’s appreciation of phonology, word meaning, and grammatical relations, as well as a speaker’s tone of voice, preceding context, and the speaker’s mood. In simplistic terms, the right hemisphere’s contributions to language and communication are typically layered on top of the foundation provided by the left hemisphere. Even weak claims for the right hemisphere’s role are extremely important clinically because injury to the right hemisphere often results in impairments that significantly reduce a patient’s ability to communicate effectively in natural settings.

Another general point concerns localization of function. Often, a group of patients with right hemisphere lesions is compared with a group of non-brain-injured controls. Although this type of comparison does not allow localization of a particular function to the right hemisphere, it still supports the weaker interpretation mentioned earlier. In addition, some studies provide strong support for right hemisphere localization by (1) directly comparing the effects of unilateral lesions of the right and left hemispheres, (2) using lateralized presentation to intact left or right hemispheres, or (3) using functional imaging (PET, fMRI) to examine “on-line” brain activation in non-brain-injured adults. There is growing support for the right hemisphere’s unique contribution to language and communication.

Prosody. Prosody refers to variation in frequency, amplitude, duration, timbre, and rhythm. Prosodic contour can be used to convey linguistic distinctions, such as distinguishing between meanings of words or phrases (“yellow jacket” meaning a kind of bee or a brightly colored piece of clothing) and between speech acts (a question versus statement signaled by a rising pitch toward the end of an utterance). Research with right hemisphere-injured patients suggests that both expressive and receptive deficits can occur, although there is disagreement across studies. In terms of production, there is some loss of control, sometimes manifested as an increased variability in pitch (specifically, fundamental frequency) after temporal and rolandic area lesions (e.g., Colsher, Cooper, and Graff-Radford, 1987). Patients with right hemisphere lesions are also impaired on a variety of discrimination and production tasks (Behrens, 1989; Weintraub, Mesulam, and Kramer, 1981).

Prosody can also be used to convey a range of emotions, such as anger or sadness. Ross (1981) has proposed a taxonomy of aprosodias to mirror the classical taxonomy of aphasias: a motor aprosodia associated with right frontal lesions, a receptive aprosodia associated with right temporal lesions, and a global aprosodia associated with extensive frontal-temporal-parietal lesions. Other research has confirmed the separation.
of an affective deficit from a linguistic prosody comprehension deficit based on direct comparison between the effects of left- and right-sided lesions (Heilman et al., 1984; Pell, 1998).

**Lexical and Phrasal Metaphor.** Several studies have used a sentence-picture matching task. Patients with right hemisphere lesions, more so than aphasic patients with left-sided lesions, tend to be overly literal, thereby missing the conventional meaning of phrasal metaphors such as “he has a heavy heart” or idioms such as “turning over a new leaf” (Van Lancker and Kempler, 1987). The characteristic literalness has been extended to single-word stimuli such as “warm,” “cold,” “deep,” and “shallow” in a semantic similarity judgment task presented to left- and right-lesioned patient groups (Brownell et al., 1984). Functional imaging in normal adults confirms that regions in the right hemisphere (including the dorsolateral prefrontal cortex, middle temporal gyrus, and the precuneus in the medial parietal lobe) are differentially activated during metaphor processing compared to literal sentence processing (Bottini et al., 1994).

Studies of lexical semantic processing by the left and right hemispheres of normal adults, together with work on discourse, have been used to support a comprehensive model. Beeman (1998), as well as others, suggests that the left hemisphere is designed for focused processing of the closest (literal) associations to a word demanded by preceding context and for actively dampening activation of alternative meanings. The right hemisphere, in contrast, is sensitive to looser, more remote associations and allows them to persist over time. Extrapolating Beeman’s model, one can imagine what happens when a potential metaphor is presented: “Microsoft Corporation is the tiger of the software industry.” The right hemisphere maintains the various associations emanating from “Microsoft,” the topic of the metaphor, and “tiger,” the vehicle; the overlapping or shared associations between “Microsoft” and “tiger” provide the ground for the metaphor.

**Discourse.** Discourse processing requires that a listener integrate meaning across sentences and from non- or paralinguistic sources to achieve an understanding of an entire story, joke, or conversation. Several studies have documented that right hemisphere lesions more than left hemisphere lesions result in decreased humor appreciation (Shammi and Stuss, 1999). While patients with right-sided lesions have no trouble appreciating that short story jokes require an incongruous punch line, they are deficient in apprehending exactly how a punch line fits with the body of a joke on a deeper level, and they have analogous problems with other types of discourse for which comprehension requires a reinterpretation (Brownell et al., 2000). Patients with right hemisphere injuries have trouble extracting gist from extended narrative even in the absence of an obvious need for reinterpretation (Hough, 1990), although there are highly constrained situations in which they are able to perform inferences that span sentence boundaries (Leonard, Waters, and Caplan, 1997). Another realm of impairment centers on nonliteral language. Several studies of sarcasm and irony comprehension suggest a problem using context (mood, sentence prosody) as a guide to uncovering a speaker’s intended meaning (Tomkins and Flowers, 1985). Similarly, a host of studies show that right-sided lesions alter patient’s production and comprehension of indirect requests, which also require consideration of the preceding context (Stemmer, Giroix, and Joannette, 1994).

An overlapping body of work examines whether an underlying social cognitive impairment affects discourse performance. The ability to explain behavior in terms of other people’s mental states, referred to as theory of mind, has been examined in several populations, including people with autism and stroke patients. Comprehension of stories and cartoons that rely on theory of mind are relatively difficult for patients with right-sided lesions, but not for aphasic patients with left hemisphere lesions (Happe, Brownell, and Winner, 1999). Also, functional imaging studies in normal adults suggest greater activation linked to theory of mind in a variety of regions, including the right middle frontal gyrus and precuneus (Gallagher et al., 2000).

There are, of course, unresolved issues. The range of language and communication skills associated with right hemisphere injury is extensive. These skills seem to represent several domains that will need to be examined separately, even though powerful unifying constructs have been explored, such as coherence (Benowitz, Moya, and Levine, 1990) and working memory (Tomkins et al., 1994). Finally, our understanding of localization of function involving the right hemisphere is poorly developed. Functions often associated with the right hemisphere may be as appropriately tied to prefrontal regions in either hemisphere (McDonald, 1993; Stuss, Gallup, and Alexander, 2001).

See also discourse; discourse impairments.

—Hiram Brownell

**References**


Right Hemisphere Language Disorders

For nearly 150 years, the language-dominant left cerebral hemisphere has dominated research and clinical concern about language disorders that accompany brain damage in adults. However, it is now well established that unilateral right hemisphere brain damage can substantially impair language and communication. The language deficits associated with right hemisphere damage, while often quite socially handicapping (Tompkins et al., 1998), are little understood. This article focuses on disorders characterized by damage restricted to the right cerebral hemisphere in adults. Such individuals may have difficulties with some basic language tasks but are not generally considered to have aphasia, because phonology, morphology, syntax, and many aspects of semantics are largely intact. About 50% of adults with right hemisphere damage have a verbal communication disorder (Joanette et al., 1990). In one study, 93% of 123 adults with right hemisphere damage in a rehabilitation center had at least one cognitive deficit with the potential to disrupt communication and social interaction (Blake et al., 2002).

Heterogeneity typifies the population of adults with right hemisphere damage: not all will have all characteristic communicative problems, and some will have no discernible problems. This heterogeneity often is unaccounted for in sample selection or data analysis, and its potential effects are compounded by the small samples in most studies of language in patients with right hemisphere damage. A related difficulty involves control group composition in research on language deficits associated with right hemisphere damage. Non-brain-damaged samples typically comprise individuals who do not have complications associated with being a patient. Individuals with left brain damage often are excluded because they cannot perform the more complex tasks that are most revealing of language functioning after...
right hemisphere damage, and because differences in impairment profiles make it difficult to equate groups for severity. Consequently, it is impossible to determine whether observed deficits are specific to right hemisphere damage. Another major issue is the lack of consensus on how to define or even what to call language deficits associated with right hemisphere damage (cf. Myers, 1999), either in totality or as individual components of an aggregate syndrome. Conceptual and terminological imprecision, and apparent overlap, are common in referring to targets of inquiry such as nonliteral language processing, inferencing, integration, and reasoning from a theory of mind (Blake et al., 2002). Conclusions about language deficits after right hemisphere damage also are complicated by intra-individual performance variability, whether due to factors such as differential task processing requirements (e.g., Tompkins and Lehman, 1998), or to time following onset of injury (Colsher et al., 1987). Finally, many language difficulties ostensibly related to right hemisphere damage stem from or are exacerbated by other perceptual and cognitive impairments, some of which are as yet unidentified but others of which have not been evaluated consistently. Chief among these complications are hemispatial neglect, other attentional difficulties, and impairments of working memory and related processing resources.

Overview and Potential Accounts of Symptoms in Characteristic Deficit Domains

Prosody. Prosody is not uniquely a right hemisphere function (Baum and Pell, 1999). However, many adults with right hemisphere damage have difficulty producing and comprehending prosody, whether it serves linguistic functions or conveys nuance and affect (Joanette, Goulet, and Hannequin, 1990; Baum and Pell, 1999). Evidence is mixed on the occurrence and nature of prosodic problems in right hemisphere damage. Speech prosody is most commonly described as flat, but by contrast may be characterized by an abnormally high fundamental frequency and variability (Colsher, Cooper, and Graff-Radford, 1987). Dysarthrias often may be a source of spoken prosodic impairment (Wertz et al., 1998). Prosodic interpretation difficulties largely may be due to perceptual deficits, apart from the linguistic or emotional characteristics of a message (Tompkins and Flowers, 1985; Joanette, Goulet, and Hannequin, 1990). Some prosodic production and comprehension difficulties also may stem from more general emotional processing deficits (Van Lancker and Pachana, 1998).

Prosodic expression and comprehension deficits can be dissociated, and Ross (1981) proposed a taxonomy for prosodic impairments, relating them to right hemisphere lesion site. However, the proposed functional-anatomical correlations have not been substantiated in other research (e.g., Wertz et al., 1998), and the neurological correlates of prosodic disruption are more complex than Ross’s framework suggests (Baum and Pell, 1999; Joanette, Goulet, and Hannequin, 1990).

Related Emotional and Nonverbal Processing Deficits. Emotional processing deficits also are a hallmark of right hemisphere damage (Van Lancker and Pachana, 1998), potentially contributing to difficulties with social exchange. Adults with right frontal lesions may be emotionally disinhibited, while those with more posterior damage may minimize and rationalize their deficits. Many adults with right hemisphere damage demonstrate reduced nonverbal animation and coverbal behaviors (Blonder et al., 1993). Some exhibit emotional interpretation deficits across modalities (pictures, body language, facial and vocal expressions, and complex discourse). Hypoarousal can occur in the presence of emotional stimuli, though not necessarily with impaired emotional recognition (Zoccolotti, Scabini, and Violani, 1982). Some work suggests problems in the way adults with right hemisphere damage apply a relatively intact appreciation of emotional material. For example, individuals with right hemisphere damage may do well inferring the affect conveyed by sentences that describe emotional situations (Tompkins and Flowers, 1985) but falter when required to match emotional infences with specific stimulus representations or settings (Cicone, Wapner, and Gardner, 1980). Emotional misinterpretation also may reflect problems appreciating the visuospatial and acoustic/prosodic stimuli in which emotional messages are embedded.

Lexical-Semantic Processing. Right hemisphere damage is not usually considered to impair lexical structure, but it has been shown to diminish performance on tasks that involve lexical-semantic processing, such as picture naming, word-picture matching, word generation, and semantic judgment. These findings have been taken to indicate a general, subtle, but specific deficit in lexical-semantic processing (e.g., Gainotti et al., 1983). However, visual-perceptual problems could account for difficulties in many such tasks. Additionally, semantic priming studies indicate no lexical-semantic processing deficit after right hemisphere damage, under either automatic or controlled activation (Joanette and Goulet, 1998; Tompkins, Fassbinder, et al., 2002). This suggests that right hemisphere damage does not affect representation and initial activation processes in the lexical-semantic system (although currently it is not possible to rule out slowed initial processing). Because clear difficulties are evident only on metalinguistic tasks that obscure lexical-semantic operations per se, the “lexical-semantic deficits” of right hemisphere damage could reflect difficulties with other specific task requirements or more general attentional or working memory limitations (Joanette and Goulet, 1998; Tompkins, Fassbinder, et al., 2002).

Adults with right hemisphere damage also can have difficulty making semantic judgments about words with metaphorical or emotional content. Thus, some investigators have suggested that right hemisphere damage impairs specific semantic domains. A prominent account of difficulties with metaphorical meanings was derived from studies of hemispheric differences in non-brain-
damaged individuals. This work suggests that as words are processed for comprehension, the right hemisphere is solely responsible for maintaining activation of a rather diffuse network of peripheral or secondary interpretations and remote associates of those words. This broad-based lexical-semantic activation is proposed to underpin figurative language interpretation, among other comprehension abilities (e.g., Beeman, 1998). By extrapolation, right hemisphere damage is assumed to impair the maintenance of weak associates and secondary meanings, making, for example, metaphorical interpretations unavailable. Contrary to this proposal, however, adults with right hemisphere lesions evince both initial, automatic priming and prolonged priming of metaphorical meanings of words (i.e., 1000 ms after hearing the word; Tompkins, 1990). Again, difficulties do not occur on implicit tasks with minimal strategic processing demands; thus, factors related to processing capacity and processing load cannot be excluded as the source of domain-specific deficits for metaphorical words in adults with right hemisphere damage (Joanette and Goulet, 1998; Tompkins, Fassbinder, et al., 2002). The processing of emotional words has not been investigated with implicit methods, so the influences of processing demands cannot be evaluated at present.

**Discourse and Conversation.** A growing literature suggests possible impairments of building, extracting, applying, or manipulating the mental structures that guide discourse processing after right hemisphere damage (Beeman, 1998; Tompkins, Fassbinder, et al., 2002). Again, contrasting findings abound (Tompkins, 1995). For example, although the output of adults with right hemisphere damage most often is described as verbose, digressive, and lacking in informative content, some produce a paucity of spoken discourse (Myers, 1999). Deficits in organizing and integrating elements of discourse structure may be evident as well. As with the lexical-semantic investigations, confounds introduced by typically metalinguistic assessment tasks (Tompkins and Baumgaertner, 1998) create difficulties for interpreting much of the literature on discourse in adults with right hemisphere lesions.

Discourse production often is investigated in terms of conversational pragmatics. In the few available studies, heterogeneity again is the rule, with some participants with right hemisphere damage displaying no deficits. Adults with right hemisphere damage most often are reported to have difficulties with eye contact and some, but not all, turn-taking parameters (Prutting and Kirchner, 1987; Kennedy et al., 1994). Idiosyncratic and ambiguous reference, poor re-use of common referents, and excessive attention to peripheral details also may occur (Chantraine, Joanette, and Ska, 1998).

Discourse comprehension problems are especially evident when adults with right hemisphere damage must reconcile multiple, seemingly incongruent inferences. They have particular problems with messages that induce conflicting interpretations, such those that contain ambiguities or violate canonical expectations. As summarized by Tompkins, Fassbinder, et al. (2002), inference generation per se is not a primary interpretive roadblock. Rather, difficulties appear to involve integration processes that are needed to revise or repair erroneous interpretations that were activated by such messages. Relatedly, adults with right hemisphere damage have difficulty synthesizing their mental representations of stimulus elements and stimulus contexts in order to determine nonliteral intent, as expressed in jokes, idioms, indirect requests, connotative meanings of words, and conversational irony. Again, right hemisphere damage does not clearly affect the representation or activation of such nonliteral meanings, and adults with right-sided lesions can represent relevant elements of stimulus contexts in nonliteral processing tasks (see Tompkins, Fassbinder, et al., 2002). Finally, adults with right hemisphere damage may perform poorly on tasks that require reasoning from a “theory of mind,” which involves an understanding of the ways in which knowledge, beliefs, and motivations guide behavior. Once again, these difficulties cannot be attributed to a failure to understand or represent individual elements of comprehension scenarios (Winner et al., 1998). Overall, for these deficit areas, impaired performance by adults with right-sided lesions is evident in conditions of relatively high processing demand.

There are several emerging accounts of difficulties experienced by adults with right hemisphere damage in constructing coherent, integrated mental structures that support discourse production and comprehension. Brownell and Martino (1998) implicate problems with “self-directed inference” (p. 325), which refers to comprehenders’ efforts to discover and elaborate an interpretive framework when overlearned interpretive routines are inadequate. Reasoning from a theory of mind also is gaining popularity as an explanatory construct (Brownell and Martino, 1998). Another prominent hypothesis derives from characterization of the normal right hemisphere’s “coarse semantic coding” properties (Beeman, 1998, p. 255), or the broad-based activation of diffuse, peripheral, and secondary meanings of words. Damage to the right hemisphere presumably creates difficulty in activating and/or maintaining the distant associates and subordinate meanings on which connotative interpretations, various inferences, and some discourse integration processes rely. However, Tompkins and colleagues (2000, 2001) demonstrated that adults with right hemisphere damage activate multiple meanings of lexical and inferential ambiguities, and that abnormally prolonged activation of contextually incompatible interpretations predicts aspects of discourse comprehension after right hemisphere lesions. These authors propose that right hemisphere damage may impair a comprehension mechanism by which contextually incompatible interpretations are suppressed, and they argue that this “suppression deficit” account accommodates a variety of other existing data. The suppression deficit and maintenance deficit views may be reconcilable by considering within-hemisphere site of lesion, a possibility currently under investigation (Tomp-
kins, 2002). More generally, Stemmer and Joanette (1998) suggest that discourse deficits in individuals with right-sided lesions reflect difficulty with constructing and integrating new conceptual models. In a different vein, Brownell and Martino (1998) maintain that many impairments associated with right hemisphere damage stem from a social disconnection or diminished interest in people. Finally, because the expression of deficits in various domains seems to be moderated by processing abilities and demands, factors related to processing capacity and processing load need to be considered in a full account of impairments and skills in individuals with right hemisphere damage (Tompkins, Blake, et al., 2002; Tompkins, Fassbinder, et al., 2002).

Impact and Management

The cognitive and behavioral problems of adults with right hemisphere lesions can interfere with judgment and social skills, family relationships, functional living activities, and the potential to return to productive work (Klonoff et al., 1990; Tompkins et al., 1998). Clinical management typically is symptom driven, although various authors emphasize the value of a theoretically oriented approach (Tompkins, 1995; Myers, 1999). Treatment that focuses on the remediation of deficits may miss the bigger picture, in which therapeutic benefit is assessed in terms of its effects on daily life activities and psychosocial functioning (Tompkins et al., 1998; Tompkins, Fassbinder, et al., 2002). Treatment research is urgently needed for this population.

See also DISCOURSE IMPAIRMENTS; PROSODIC DEFICITS.

— Connie A. Tompkins and Wiltrud Fassbinder

References


Tompkins, C. A. (2002). Comprehension impairment and right brain damage. Grant funded by the National Institutes of Health: National Institute on Deafness and Other Communication Disorders (DC01820).


Tompkins, C. A., Fassbinder, W., Lehman-Blake, M. T., and Baumgaertner, A. (2002). The nature and implications of...


Further Readings


Segmentation of Spoken Language by Normal Adult Listeners

Listening to spoken language usually seems effortless, but the processes involved are complex. A continuous acoustic signal must be translated into meaning so that the listener can understand the speaker’s intent. The mapping of sound to meaning proceeds via the lexicon—our store of known words. Any utterance we hear may be novel to us, but the words it contains are familiar, and
to understand the utterance we must therefore identify the words of which it is composed.

We know a great many words; an educated adult’s vocabulary has been estimated at around 150,000 words. Entries in the mental lexicon may include, besides standalone words, grammatical morphemes such as prefixes and suffixes and multiword phrases such as idioms and cliches. Languages also differ widely in how they construct word forms, and this too will affect what is stored in the lexicon. But in any language, listening involves mapping the acoustic signal onto stored meanings.

The continuity of utterances means that boundaries between individual words in speech are not overtly marked. Speakers do not pause between words but run them into one another. The problem of segmenting a speech signal into words is compounded by the fact that words themselves are not highly distinctive. All the words we know are constructed of just a handful of different sounds; on average, the phonetic repertoire of a language contains 30–40 contrasting sounds (Maddieson, 1984). As a consequence, words inevitably resemble one another without a break (stray, strain, train, range, jack, and so on, as well as the intended words). The only segmentation that is logically required is to differentiate cues to adjacent phonemes to overlap in time are especially prominent. The presence of a word boundary. For instance, in English the phoneme sequence [mg] never occurs word-initially, so the occurrence of this sequence implies at least a syllable boundary (e.g., to narrow the field of candidates to only robe and rope, eliminating rogue, road, and wrote).

Evidence for continuous evaluation comes from experiments in which listeners perform lexical decision (judging whether a spoken string is indeed a real word) on speech that has been cross-spliced so that the coarticulatory effects are no longer reliable. Thus, when listeners hear trook they should respond “no”—trook is not a word. If trook is cross-spliced so that a final -t is appended to a troo- from either trook or troop (which give coarticulatory cues to an upcoming velar or bilabial consonant, respectively), then responses are slower than if the cues match. This shows that listeners are sensitive to the coarticulatory mismatch and must have processed the consonant place cues in the vowel. However, the responses are still slower when the mismatching troo- comes from troop than when it comes from trook. This suggests that the processing of consonant cues in the vowel has caused activation of the existing compatible real-word troop (Marslen-Wilson and Warren, 1994; McQueen, Norris, and Cutler, 1999).

Second, multiple candidate words are simultaneously activated during the listening process, including words that are merely accidentally present in a speech signal. Thus, hearing strange-acting may activate stray, train, range, jack, and so on, as well as the intended words.

Evidence for multiple activation comes from cross-modal priming experiments in which a word-initial fragment facilitates recognition of different words that it might become. Thus, lexical decision responses for visually presented “captain” or “captive” are both facilitated when listeners have just heard the fragment capt- (compared with some other control fragment). Moreover, both are facilitated even if only one of them matches the context (Zwitserlood, 1989).

Third, there is active competition between alternative candidate words. The more active a candidate word is, the more it may suppress its rivals, and the more competitors a word has, the more suppression it may undergo. Evidence for competition between simultaneously activated candidate words comes from experiments in which listeners must spot any real words occurring in spoken nonsense strings. If the rest of the string partially activates a competitor word, then spotting the real embedded word is slowed. For instance, listeners spot mess less rapidly in domess (which partially activates domestic, a competitor for the same portion of the signal that supports mess) than in nemess (which supports no other word; McQueen, Norris, and Cutler, 1994; see also Norris, McQueen, and Cutler, 1995; Vroomen and de Gelder, 1995; Soto-Faraco, Sebastián-Gallés, and Cutler, 2001).

Because activated and competing words need not be aligned with one another, the competition process offers a potential means of segmenting the utterance. Thus, although recognition of strange-acting may involve competition from stray, range, jack, and so on, this will eventually yield to joint inhibition from the two intended words, which receive greater support from the signal.

Adult listeners can also use information which their linguistic experience suggests to be correlated with the presence of a word boundary. For instance, in English the phoneme sequence [mg] never occurs word-internally, so the occurrence of this sequence must imply a word boundary (some go, tame goose); sequences such as [pf] or [ml] or [zw] never occur syllable-internally, so this sequence implies at least a syllable boundary (cupful, seemly, beeswax). Listeners more rapidly spot embedded words whose edges are aligned with such a boundary-correlated sequence (e.g., rock is spotted more easily in foomrock than in foogrock; McQueen, 1998). Also,
words that begin with a common phoneme sequence are easier to extract from a preceding context than words that begin with an infrequent sequence (e.g., in golnook versus golnag, it will be easier to spot nag, which shares its beginning with natural, navigate, narrow, nap, and many other words; van der Lugt, 2001; see also Cairns et al., 1997).

These latter sources of information are, of course, necessarily language-specific. It is a characteristic of a particular vocabulary that more words begin with the na- of nag than with the noo- of nook; likewise, it is vocabulary-specific that sequences such as [pf] or [zw] or [ml] cannot occur within a syllable. Each of these three sequences is in fact legitimately syllable-internal in some language ([pf], for instance, in German: Pferd, Kopf).

Other language-specific information is also used in segmentation, notably rhythmic structure. In languages such as English and Dutch, most words begin with stressed syllables, and listeners find it easier to segment speech at the onset of stressed syllables (Cutler and Norris, 1988; Vroomen, van Zon, and de Gelder, 1996). This can be clearly seen in segmentation errors, as when a pop song line She’s a muscular boy— the strong syllable void is taken to be the onset of a new word, while the weak syllables to and a- are taken to be noninitial (Cutler and Butterfield, 1992).

The stress rhythm of English and Dutch is not universal; many other languages have different rhythmic structures. Indeed, syllabically based rhythm in French is accompanied by syllabic segmentation in French listening experiments (Meehler et al., 1981; Cutler et al., 1986; Kolinsky, Morais, and Cluytens, 1995), while moraic rhythm in Japanese likewise accompanies moraic segmentation by Japanese listeners (Otake et al., 1993; Cutler and Otake, 1994).

Thus, although the type of rhythm is language-specific, its use in speech segmentation seems universal. Other universal constraints on segmentation exist, for example, to limit activation of spurious embedded competitors. It is harder to spot a word if the residual context contains only consonants (thus, apple is harder to find in fapple than in vuffapple; Norris et al., 1997), an effect explained as a primitive filter selecting for possible words— vuff is not a word, but it might have been one, while f could never be a word. This constraint would operate to rule out many spuriously present words in speech (such as tray and ray in stray). It is not affected by what may be a word in a particular language (Norris et al., 2001; Cutler, Demuth, and McQueen, 2002) and thus appears to be universal.

The ability to extract words from continuous speech starts early in life, as shown by experiments in which infants listen longer to passages containing words that they had previously heard in isolation than to wholly new passages (Jusczyk and Aslin, 1995); none of the passages can be comprehended by these young listeners, but they can recognize familiar strings embedded in the fluent speech. One-year-olds also detect familiar strings less easily if they are embedded in a context without a vowel (e.g., rest is found less easily in crest than in caressed; Johnson et al., 2003); that is, they are already sensitive to the apparently universal constraint on possible words.

Finally, segmentation of second languages in later life is not aided by the efficiency with which listeners exploit language-specific structure in recognizing speech. Segmentation procedures suitable for the native language can be inappropriately applied to non-native input (Cutler et al., 1986; Otake et al., 1993; Cutler and Otake, 1994; Weber, 2001). This is one effect making listening to a second language paradoxically harder than, for instance, reading the same language.

See also PHONOLOGY AND ADULT APHASIA.

—Anne Cutler

References


The term **semantics** refers to linguistic meaning. Semantic development involves mapping linguistic forms onto our mental models of the world and organizing these maps into networks of related information. Semantic processing involves the comprehension and production of meaning as conveyed by linguistic forms, their combinations, and the nonlinguistic contexts in which they occur.

In this article, semantic development in selected populations of children with language disorders is described via intralinguistic referencing. Included are summaries of deficits in semantic processing that characterize each population as a whole, at early and later points in development.

### Early Semantic Development in Children with Developmental Language Disorders

Children with autism or other pervasive developmental disorders (PDDs) typically demonstrate semantic systems that are weak relative to the formal systems of syntax, morphology, and phonology. This weakness is manifested as use of words without regard to conventional meaning, context-bound extensions of word meaning, and confusion regarding the mapping of personal pronouns onto their referents. Social and cognitive deficits are thought to contribute to this weakness. Whereas normally developing children readily make inferences about word meanings by reading social and contextual cues, such as the speaker’s eye gaze and intentions, children with autism do not (Baron-Cohen, Baldwin, and Crowson, 1997). This failure is often viewed as part of a broader deficit in theory of mind. The theory of mind deficit is further reflected in a particular limitation of children with autism in the acquisition of words for cognitive states (Tager-Flusberg, 1992).

In **specific language impairment** (SLI), semantics is generally a strength relative to the formal domains of **morphosyntax and syntax**. However, children with SLI do demonstrate some semantic delay relative to their normal age-mates. For example, the appearance of first meaningful words is delayed an average of 11 months (Trauner et al., 1995). This delay is maintained throughout the early preschool years and is measurable with receptive vocabulary tests and experimental word-learning paradigms, which show that children with SLI are poor at fast mapping and long-term retention of words (Rice et al., 1994). Even when these children succeed at word mapping, the semantic elaboration of those words in the lexicon is sparser than age expectations would predict. This sparseness is associated with greater numbers of naming errors during picture naming (McGregor et al., 2002). In their spontaneous speech, these children exhibit less lexical diversity, especially verb diversity, than their normal age-mates (Watkins et al., 1995). Finally, for children with SLI, processing the subtle meanings carried by grammatical morphemes, derivational morphemes, and word order is particularly problematic (Bishop, 1997).

Short-term (working) memory deficits likely contribute to the delayed semantic development of children with SLI (Gathercole and Baddeley, 1990). These deficits are associated with slower word learning, presumably because features of new words or their referents are not sufficiently represented in short-term memory to be committed to the long-term lexical store. Another
possible contributor to weak lexical semantics is the limited ability of children with SLI to benefit from syntactic bootstrapping. On tasks requiring the acting out of unfamiliar verbs, children with SLI do not infer meanings from syntactic bootstrapping as well as their younger language-matched controls (van der Lely, 1994), or they require a great deal of processing effort to do so (O’Hara and Johnston, 1997).

Down syndrome is another condition in which intralinguistic referencing reveals semantics as a relative strength. In the Down syndrome population, receptive single-word vocabulary is unaffected in early childhood and may exceed nonverbal cognitive ability by adolescence (Chapman, 1995). However, children with Down syndrome demonstrate semantic delays relative to their normal (mental-age-matched) peers. Soon after the onset of first words, expressive vocabulary development begins to lag (Cardoso-Martins, Mervis, and Mervis, 1985). As children with Down syndrome learn to combine words into sentences, their expressive delays in lexical semantics relative to their MLU-matched peers are manifested as a lower rate of verbal use per sentence (Hesketh and Chapman, 1998). As in SLI, children with Down syndrome have difficulty expressing meanings with grammatical morphemes and noncanonical word order (Kumin, Councill, and Goodman, 1998). Limited diversity of vocabulary is characteristic of their discourse (Chapman, 1995). Also, similar to children with autism, these children have difficulty expressing internal states; use of words to express volition, ability, or cognition is particularly compromised (Beeghly and Cicchetti, 1997).

The semantic delays of children with Down syndrome are multidetermined. Degree of mental retardation limits development of the semantic-conceptual system. Fluctuating mild to moderate hearing loss, which is highly prevalent in this population, also has some effect (see OTITIS MEDIA: EFFECTS ON CHILDREN’S LANGUAGE). Finally, as in SLI, limited short-term memory is thought to play a role (Chapman, 1995).

**Early Semantic Development in Children with Acquired Language Disorders**

In children who have acquired language disorders subsequent to unilateral focal lesions, intralinguistic profiles often vary according to location of the lesion in the brain. Children with right hemisphere lesions tend to have more pronounced deficits in semantics than in formal aspects of language; children with left hemisphere lesions demonstrate the reverse pattern (Eisele and Aram, 1994). This generalization is gross, and semantic involvement will vary across individuals. Overall, compared to their normal age-mates, children with unilateral lesions present with late onset of first words in both comprehension and production, with right-lesioned children having more significant comprehension deficits than left-lesioned children (Thal et al., 1991). In experimental word-learning studies, unilateral brain-lesioned children require more teaching trials than their normal peers to demonstrate comprehension (Keefe, Feldman, and Holland, 1989). Also, these children have difficulty building semantic networks. For example, when asked to select two antonyms, synonyms, or class coordinates from lists of four words, both right- and left-lesioned schoolchildren made significantly more errors than their normal age-mates (Eisele and Aram, 1995). In spontaneous speech, their semantic deficits are manifested as low numbers and diversity of words (Feldman et al., 1992).

Type and extent of insult, as well as age at onset, influence semantic processing in children with acquired language disorders. In general, cognitive disorganization and poor long-term memory exacerbate semantic deficits in this population (Levin et al., 1982).

**Later Semantic Development**

Semantic weaknesses in older children and adolescents extend to the discourse level. In order to process discourse, children must integrate meanings across sentences (linguistic context) with regard to shared information, communicative goals, and the physical setting (nonlinguistic context). Not surprisingly, most individuals with language disorders, no matter their diagnostic category, have significant difficulties with semantics at the discourse level. Children with acquired language disorders have trouble recalling meaningful propositions from stories and, when producing stories, they have difficulty conjoining meaning across sentences via referential and lexical ties (Ewing-Cobbs et al., 1998). Children with developmental language disorders present with similar problems (Bishop and Adams, 1992). Processing the implicit meanings, abstract meanings, and nonliteral meanings characteristic of sophisticated discourse is also problematic for many children with either acquired or developmental language disorders. Some argue that a subgroup of children with developmental language disorders, who share certain characteristics of both SLI and PDD syndromes, have semantic-pragmatic disorders as their primary area of deficit. Such children are particularly deficient in all higher level semantic attainments (Rapin and Allen, 1983).

**Semantic Development and Reading**

Children with language impairments of both developmental and acquired types have a higher incidence of reading impairment than the general population (see LANGUAGE IMPAIRMENT AND READING DISABILITY). For example, roughly half of a group identified as having SLI at 5 years were reading-impaired at 15 years (Stothard et al., 1998). Part of this reading difficulty relates to weaknesses in semantics (poor phonological and grammatical processing also plays a role). These children may not have sufficient semantic representations on which to map orthography. Furthermore, they are less able to use semantic bootstrapping to infer the meaning of words in context (Snowling, 2000).

The developmental relation between lexical semantics and reading is reciprocal. Poor knowledge in the seman-
tic domain contributes to poor reading, and poor reading exacerbates lags in semantic development. Children who have poor reading skills read fewer words over the course of a year and are less able to learn the words they do read than those who have good reading skills (Nagy and Anderson, 1984).

Children with hyperlexia, often viewed as a subgroup of the PDD spectrum, demonstrate a preoccupation with decoding written words to the exclusion of meaning. Word recognition skills in these children are reported to be as much as 7 years in advance of grade-level expectations, whereas reading comprehension is at or below grade level (Whitehouse and Harris, 1984). Such children demonstrate that form and meaning can be sharply divorced in developmental disorders.

**Future Trends**

Recent years brought improved methods for identifying lexical semantic deficits in children. Newly developed parent-report inventories provide valid estimates of the size of receptive lexicons in children functioning below the 16-month level and expressive lexicons in children functioning below the 30-month level (Fenson et al., 1993). For more skilled children, new ways of quantifying lexical-semantic abilities from discourse samples provide valid diagnostic indicators (Miller, 1996).

The near future promises a burgeoning interest in semantic development in children with language disorders. Increasingly, investigators, motivated by social interactionist and dynamical systems theories, are demonstrating that semantic ability is not a static collection of knowledge but a system that emerges from interactions among and between knowledge, context, and processing demands in real time. Aiding investigation of real-time semantic processing are new technologies and methods such as semantic priming, eye tracking, and event-related potentials. Aiding identification of semantic deficits is the inclusion of dynamic word-learning tasks in diagnostic batteries. When these sophisticated methods are widely employed, both the quality and the quantity of semantic representations in children affected by language impairments will receive increased attention.

—Karla M. McGregor

**References**


Further Readings


Social Development and Language Impairment

Social skills determine to a large extent the success we enjoy vocationally and avocationally and the amount of satisfaction we derive from our personal relationships. Social skill deficits in childhood have long been associated with a variety of negative outcomes, including criminality, underemployment, and psychopathology (cf. Gilbert and Connolly, 1991). The potential impact of developmental language impairments on social development is one of the most important issues facing families, educators, and speech-language pathologists. Unfortunately, there is little agreement on the definition of social skills as a psychological construct, making service planning in this area challenging. In their review, Merrell and Gimpel (1998) presented no less than 16 different definitions that enjoy wide currency and reflect the interests of a variety of disciplines, including psychology, psychiatry, special education, and social work. However, commonalities across these different perspectives can be extracted. In a meta-analysis of 21 multivariate studies that classified children’s social skills (total N = 22,000), Caldarella and Merrell (1997) identified five core...
dimensions: peer relations, self-management, academics, compliance, and assertion.

What are the interrelationships between language impairments and these important areas of social development? Language impairments occur with a large variety of developmental disorders, and some, such as mental retardation, autism, and pervasive developmental delay, include social skill deficits as a primary diagnostic feature. In order to address the question of how language impairment uniquely affects social development, however, we need to examine the social skills of children with specific language impairment (SLI) (see Specific Language Impairment in Children). SLI refers to a language deficiency that occurs in the absence of other conditions commonly associated with language disorders in children. Children with SLI show normal hearing, age-appropriate scores on nonverbal tests of intelligence, and no obvious signs of neurological or socioemotional impairment. Children with SLI represent a heterogeneous group, and significant individual differences exist among children diagnosed with this disorder. However, a common profile in young, English-speaking children with SLI is a mild to moderate deficit in a range of language areas and a more significant deficit in the use of grammatical morphology.

According to Caldarella and Merrell (1997), a large number of social skills contribute to the dimension peer relations. These skills include specific discourse/pragmatic behaviors such as complimenting others and inviting others to play, as well as more general social attributes such as peer acceptance. Several studies have examined the peer interactions of preschool and school-age children with SLI and have documented the detrimental effect that language impairments can have on this area of social development. For example, children with SLI are likely to be ignored by their typically developing peers, respond less often when their peers make initiations, and rely more on adults to mediate their interactions (Craig and Evans, 1989; Hadley and Rice, 1991; Rice, Sell, and Hadley, 1991; Craig and Washington, 1993; Brinton, Fujiki, and Higbee, 1998; Brinton, Fujiki, and McKee, 1998). Sociometric analyses confirm further the impression that children with SLI experience limited peer acceptance (Gertner, Rice, and Hadley, 1994; Fujiki, Brinton, Hart, et al., 1999). Some studies suggest that problems in peer group acceptance may extend to difficulties establishing adequate friendships (Fujiki, Brinton, Morgan, et al., 1999), whereas others have reported no differences between children with SLI and typically developing children in the number and quality of close friendships (Redmond and Rice, 1998).

Although peer relations represent the sine qua non of social development, there are other important social skills. Self-management refers to the ability to control one’s temper, follow rules and limits, and compromise with others (Caldarella and Merrell, 1997). A few studies have assessed this dimension in children with SLI. Stevens and Bliss (1995) examined conflict resolution abilities in 30 children with SLI in grades 3 through 7 and found no significant differences between this group and a group of grade-matched typically developing controls during a role-enactment activity. Children with SLI produced fewer strategies during a more verbally demanding hypothetical problem-solving activity. Brinton, Fujiki, and McKee (1998) examined the negotiation skills of six children with SLI (8 to 12 years old) during conversational interactions with typically developing peers and found that children with SLI contributed fewer and less mature negotiation strategies. The strategies used by children in the SLI group resembled those produced by a younger group of children of equivalent language levels.

The third cluster of social skills identified by Caldarella and Merrell's (1997) meta-analysis was the academics dimension. This dimension captures behaviors regarded by teachers as important to school adjustment and is represented by such skills as completing tasks independently, following teachers' directions, and producing quality work. Indirect evidence for problems in this dimension comes from studies of SLI that have used rating scales to evaluate children’s socioemotional characteristics. For example, Redmond and Rice (1998) compared standardized parent and teacher ratings of 17 children with SLI collected at kindergarten and first grade to ratings collected on typically developing children. The particular rating scales used included several items that relate to important academic skills (e.g., has difficulty following directions; fails to carry out tasks; messy work). These investigators found significant differences between groups on the teacher ratings of these problems but not on the parent ratings, suggesting that the social performance of children with SLI varies significantly across situations, depending on the verbal demands placed on them and the expectations of others. Levels of academic success may also influence academic social behaviors. In an epidemiological study of 164 second-grade children with language impairments, Tomblin et al. (2000) found that levels of classroom behavior problems were higher among children with SLI who also had reading disabilities than among children with SLI alone (see Language Disorders and Reading Disabilities).

Prosocial behaviors such as cooperation and sharing are captured by Caldarella and Merrell’s compliance dimension. Information on the consequences of SLI in this area of social skill development is limited. Farmer (2000) compared the performances of 16 10-year-old children with SLI with that of a group of typically developing children on a standardized teacher rating scale of prosocial behaviors and found no significant differences.

The final social skills dimension in Caldarella and Merrell’s taxonomy is assertion. Several studies suggest children with SLI experience particular difficulty in this area. Craig and Washington (1993) examined the conversational skills of five 7-year-old children with SLI as they attempted to access ongoing peer interactions and found that three of the five children had considerable difficulty asserting themselves in this situation. Brinton et al. (1997) replicated these results with older children.
(8–12 years old). Children with SLI have also been consistently characterized as shy, passive, and withdrawn by parents and teachers (Tallal, Dukette, and Curtiss, 1989; Fujiki, Brinton, and Todd, 1996; Redmond and Rice, 1998; Beitchman et al., 2001). Results from a recent 14-year longitudinal study of 77 children with speech and language impairments suggest that characterizations of low assertiveness may be longstanding and continue at least into young adulthood for some children with SLI (Beitchman et al., 2001).

In sum, a small but growing body of research suggests that language impairments place children at risk for negative social consequences, specifically in the areas of assertion and peer relations. Problems in these areas may be particularly detrimental for children with SLI because they can contribute to what Rice (1993) has described as a “negative social spiral.” Rice suggested that in response to repeated instances of communicative failure, children with SLI may withdraw from peer interactions or rely more on adults to mediate peer interactions. However, these behavioral adjustments may turn out to be counterproductive for both social and linguistic development because they limit children’s access to important socialization experiences and opportunities to improve their limited language skills.

Studies of children with SLI have consistently reported high levels of variability in social skill performance, a finding that has important clinical and research implications. Social skill deficits do not appear to be an inevitable consequence of developmental language impairments, nor can social skill differences between children be inferred from differences in either the type or severity of language deficits (e.g., Brinton and Fujiki, 1999; Fujiki et al., 1999; Donlan and Masters, 2000). Given the heterogeneity of children with developmental language disorders, it is important that treatment teams supplement the assessment of language impairment with a separate assessment of social skills in the areas of peer relations, self-management, academics, compliance, and assertion across different situational contexts (home, classroom, playground).

Future investigations may reveal that variability in children with language impairments is better accounted for by uncontrolled confounding factors. For example, although many studies suggest that SLI and attention-deficit/hyperactivity disorder commonly co-occur (cf. Cohen et al., 2000) the potential influence of this comorbidity on social skill development has not yet been considered. Likewise, a small portion of children diagnosed with SLI demonstrate limitations in social cognition commonly associated with autism and pervasive developmental delay. There has been a longstanding controversy over the diagnostic boundaries between SLI and autism spectrum disorders (cf. Bishop, 2000), and social skill outcomes may be an important distinguishing characteristic of children who fall outside preconceived categories. Large-scale investigations comparing the social skills of children with SLI only with those of children with SLI and other comorbid disorders are needed to delineate which language skills are associated with specific areas of social skill development. The results of this line of research will inform important areas of clinical practice, such as diagnosis, prognosis, and treatment.

See also PSYCHOSOCIAL PROBLEMS ASSOCIATED WITH COMMUNICATIVE DISORDERS.

—Sean M. Redmond

References


Social Development and Language Impairment


Further Readings


Specific Language Impairment in Children

*Specific language impairment* (SLI) is a term that is applied to children who show a significant deficit in their spoken language ability with no obvious accompanying problems such as hearing impairment, mental retardation, or neurological damage. This type of language disorder is regarded as developmental in nature because affected children exhibit language learning problems from the outset.

Although SLI is receiving increased attention in the research and clinical literature, it is not a newly discovered disorder. Children meeting the basic definition of SLI have been described in the literature since the 1800s, but have been given a wide range of clinical labels. More recent clinical labels used for children with SLI include developmental aphasia, developmental dysphasia, and developmental language disorder. The last continues to be used in the *DSM-IV* classification system, with the subtypes of “expressive” and “receptive and expressive” (American Psychiatric Association, 1994). These subtypes acknowledge that some children with SLI may have significant limitations primarily in the area of language production, whereas others may have major limitations in both the comprehension and production of language.

The prevalence of SLI is estimated to be approximately 7% among 5-year-olds, based on epidemiological data (Tomblin et al., 1997). Males outnumber females; the most recent evidence suggests a ratio of approximately 1.5:1.

Children with SLI are two to three times more likely than typically developing children to have parents or siblings with a history of language problems (Tallal, Ross, and Curtiss, 1989; Tomblin, 1989; Tallal et al., 2001). For children with family histories of language problems, there is reason to suspect a genetic basis rather than a primary environmental basis. Concordance rates for SLI are considerably higher for monozygotic twins than for same-sex dizygotic twins (Bishop, North, and Donlan, 1995). Rapid progress is being made in the genetic study of SLI. For a well-studied three-generational family that includes a high proportion of members with SLI, the evidence implicates a region on the long arm of chromosome 7 (Fisher et al., 1998). The same region was identified in a group of children with SLI who participated in an epidemiological study (Tomblin, 1999). However, other recent studies of clinically referred cases of SLI have revealed prominent areas of linkage on chromosomes 16 and 19, but not on chromosome 7 (SLI Consortium, 2002). Clearly, further refinement is needed before the genetic basis for SLI is fully understood (see Bishop, 2002).

In recent years, neuroanatomical evidence of differences between individuals with SLI and typically developing individuals has appeared in the literature (see Ahmed, Lombardino, and Leonard, 2000, for a recent review). The specific differences observed have varied across studies. For example, symmetry of the right and left perisylvian areas seems to be more likely in children with SLI than in controls. Interestingly, this pattern can also be seen in parents or siblings of children with SLI even when they do not exhibit a language disorder. Other studies have revealed a higher likelihood of atypical neuroanatomical patterns in children with SLI than in controls, but differences among the children with SLI in the particular pattern seen (e.g., ventricular enlargement, central volume loss).

Although a diagnosis of SLI is not given to children unless they meet the criteria noted above, many children with SLI nevertheless show subtle weaknesses in other areas. For example, as a group, these children are slower and less accurate on nonlinguistic cognitive tasks such as mental rotation (e.g., Miller et al., 2001), and less coordinated than their typically developing same-age peers (Powell and Bishop, 1992; Hill, 2001). These findings have led to proposals about the possible causes of SLI, but the presence of children with SLI who show none of these accompanying weaknesses raises the possibility that SLI and subtle cognitive and motor weakness are comorbid. That is, the conditions that cause SLI frequently co-occur with conditions that cause these other problems, but the latter are not responsible for SLI.

For children with SLI whose language problems are still present at 5 years of age, difficulties with language may continue into adolescence and even adulthood (Bishop and Adams, 1990; Beitchman et al., 1996). Comparisons of young adults with a history of SLI and same-age adults with no such history reveal differences favoring the latter on a range of spoken production and comprehension tasks (Tomblin, Freese, and Records, 1992).

Children with SLI are at greater risk for reading deficits than children with typical language development. This observation can be explained in part by the fact that children with SLI and those with developmental dyslexia are overlapping populations (McArthur et al., 2000). For example, prospective study of children from homes with a positive history of dyslexia reveals significantly more difficulties with spoken language than children with no such family history (Scarborough, 1990).

The language difficulties experienced by children with SLI cover most or all areas of language, including
vocabulary, morphosyntax, phonology, and pragmatics. However, these areas of language are rarely affected to the same degree. In English, vocabulary and pragmatic skills are often relative strengths, whereas phonology and especially morphosyntax are relative weaknesses. This profile is not seen in all English-speaking children meeting the criteria for SLI. For example, some children with SLI show notable word-finding problems. There have been several attempts at determining whether the differences seen among children with SLI constitute distinct subtypes or instead represent different points on a continuum. Resolution of this issue will be important, as identification of the correct phenotype of SLI will be necessary for further progress in the genetic study of this disorder.

The heterogeneity of SLI notwithstanding, certain symptoms may have the potential to serve as “clinical markers” of SLI. For English-speaking children, two measures seem especially promising. One is a measure of the children’s use of grammatical morphemes pertaining to grammatical tense and agreement, such as regular past -ed, third person singular -s, and copula and auxiliary forms of be (Rice and Wexler, 1996). The second is a measure of children’s ability to repeat nonsense words containing several syllables (e.g., Dollaghan and Campbell, 1998). Both of these measures are quite accurate in distinguishing children with SLI from their normally developing age mates.

The linguistic profile of relative strengths and weaknesses in SLI seems to be shaped to a significant degree by the language being acquired. For example, children with SLI acquiring inflectionally rich languages such as Italian and Hebrew are not as severely impaired as their English-speaking counterparts in their use of grammatical inflections pertaining to tense and agreement. On the other hand, Swedish-speaking children with SLI show more serious problems in using appropriate word order than do children with SLI acquiring English (see Leonard, 1998, for a recent review).

Evidence for the efficacy of intervention is abundant in the literature. For example, for preschool-age children with SLI, approaches such as recasting have been relatively successful (Camarata and Nelson, 1992; Fey, Cleave, and Long, 1997). However, although the gains made in intervention usually go well beyond those that can be expected by maturation alone, no intervention approach has led to dramatic and rapid language gains by these children on a consistent basis. This is especially true when gains are defined in terms of use in spontaneous speech.

Attempts to explain the nature of SLI vary considerably. Most of these accounts focus on the extraordinary grammatical deficits often seen in children with SLI. It is possible to classify these alternative accounts according to their principal assumptions. Some accounts assume that children with SLI lack particular types of grammatical knowledge. For example, the “extended optional infinitive” account assumes that children with SLI go through a protracted period during which they assume that tense and agreement are optional rather than obligatory in main clauses (Rice and Wexler, 1996). The “representational deficit for dependent relationships” account assumes that children with SLI fail to grasp that movement or checking of grammatical features is obligatory (van der Lely, 1998).

Other types of accounts assume that children with SLI might have the potential to acquire normal grammar but have limitations in processing that slow their identification and interpretation of the relevant input and their ability to retrieve this information for production. In some cases the processing limitation is assumed to be quite general (Johnston, 1994; Ellis Weismer, 1996). In other cases the limitation is assumed to be specific to particular operations, such as phonological processing (Chiat, 2001) or the processing of brief or rapidly presented auditory information (e.g., Tallal et al., 1996).

Future research on SLI will make two types of contributions. Most obviously, greater understanding of this disorder should lead to more effective methods of assessment, treatment, and prevention. In addition, because the language disorder seen in SLI may in many cases occur in the absence of accompanying impairments, it constitutes a challenge for theories of language learning to explain.

See also LANGUAGE DISORDERS IN SCHOOL-AGE CHILDREN: OVERVIEW; SPEECH DISORDERS IN CHILDREN: A PSYCHOLINGUISTIC PERSPECTIVE.

—Laurence B. Leonard

References


**Further Readings**


Syntactic Tree Pruning

Agrammatism is a syntactic deficit following damage to the left hemisphere, usually in Broca’s area and its vicinity (Zurif, 1995). The traditional view concerning speech production in agrammatism was that syntactic ability is completely lost, and agrammatic aphasics rely on nonlinguistic strategies to concatenate words into sentences (Goodglass, 1976; Berndt and Caramazza, 1980; Caplan, 1985), or that all functional elements are impaired in agrammatic speech production (Grodzinsky, 1990; Ouhalla, 1993). However, in recent years empirical evidence has accumulated to suggest that the deficit is actually finer-grained.

Speech production in agrammatism shows an intricate and intriguing pattern of deficit. Individuals with agrammatism correctly inflect verbs for agreement but substitute tense inflection; they produce well-formed yes/no questions (in some languages), but not Wh questions; they can produce untensed embedding but not full relative sentences, and coordination markers but not subordination markers. The tree pruning hypothesis (TPH) was suggested by Friedmann (1994, 2001) and Friedmann and Grodzinsky (1997) to account for these seemingly unrelated deficits and for the dissociations between spared and impaired abilities within and across languages. The TPH is a linguistic generalization, formulated within the generative grammar framework, and was suggested to account for production only. (For a syntactic account of agrammatic comprehension, see trace deletion hypothesis; Grodzinsky, 2000.)

According to the TPH, individuals with agrammatic aphasia are unable to project the syntactic tree up to its highest node, and their syntactic tree is “pruned.” As a result, syntactic structures and elements that require the high nodes of the tree are impaired, but structures and elements that involve only low nodes are preserved (Fig. 1).

According to syntactic theories within the generative tradition (e.g., Chomsky, 1995), sentences can be represented as phrase markers or syntactic trees. In these syntactic trees, content and function words are represented in different nodes (head nodes and phrasal nodes) (Fig. 1). Functional nodes include, among others, inflectional nodes: an agreement phrase (AgrP), which represents agreement between the subject and the verb in person, gender, and number, and a tense phrase (TP), representing tense inflection of the verb. Finite verbs move from V\textsuperscript{0}, their base-generated position within the VP, to Agr\textsuperscript{0} and then to T\textsuperscript{0} in order to check (or collect) their inflection. Thus, the ability to correctly inflect verbs for agreement and tense crucially depends on the AgrP and TP nodes.

The highest phrasal node in the tree is the complementizer phrase (CP), which hosts complementizers such as “that” and Wh morphemes such as “who” and “what” that moved from the base-generated position within the VP. Thus, the construction of embedded...
sentences and Wh questions depends on the CP node being intact and accessible.

Crucially, the nodes are hierarchically ordered in the syntactic tree: the lowest node is the verb phrase, the nodes above it are the agreement phrase and the tense phrase (in this order according to Pollock, 1989), and the complementizer phrase is placed at the highest point of the syntactic tree.

The TPH uses this hierarchical order and suggests that in agrammatism, the syntactic tree is pruned from a certain node and up. Persons with agrammatic aphasia who cannot project the syntactic tree up to the TP node are not able to produce structures that require TP or the node above it, CP. Persons with agrammatic aphasia whose tree is pruned at a higher point, CP, are unable to produce structures that involve the CP. Importantly, nodes below the pruning site are intact, and therefore structures that require only low nodes, such as AgrP and VP, are well-formed in agrammatic production.

How does tree pruning account for the intricate pattern of loss and sparing in agrammatic production? The implications of tree pruning for three syntactic domains—verb inflection, embedding, and question production—are examined here.

Inflections are impaired in agrammatic production, but in a selective way. In Hebrew and in Palestinian Arabic, tense inflection was found to be severely impaired, whereas agreement inflection was almost intact in a set of constrained tasks such as sentence completion, elicitation, and repetition (Friedmann, 1994, 2001; Friedmann and Grodzinsky, 1997). Subsequent studies reported a similar dissociation between tense and agreement in Spanish, Dutch, German, and English. If tense and agreement reside in different nodes in the syntactic tree, and if TP is higher than AgrP, this dissociation is explained by tree pruning: the TP is inaccessible, and therefore tense inflection is impaired, whereas the AgrP node, which is below the pruning site, remains intact, and subsequently agreement inflection is intact. This can also account for the findings from German and Dutch, according to which individuals with agrammatic aphasia frequently use nonfinite verbs in sentence-final position, instead of the required fully inflected main verbs in second position (Kolk and Heeschen, 1992; Bastiaanse and van Zonneveld, 1998). The lack of TP and CP prevents persons with agrammatic aphasia from moving the verbs to T to collect the required inflection, and to C to second position, and therefore they produce the verbs in a nonfinite form, which does not require verb movement to high nodes. Verbs that do not move to high nodes stay in their base-generated position within VP, which is the sentence-final position (see Friedmann, 2000).

Individuals with agrammatic aphasia are also known to use only simple sentences and to avoid embedded sentences. When they do try to produce an embedded sentence, either a relative clause (such as “I saw the girl that the grandmother drew”) or a sentential complement of a verb (“The girl said that the grandmother drew her”), they fail, and stop before the complementizer (the embedding marker “that,” for example), omit the complementizer, use direct instead of indirect speech, or produce an ungrammatical sentence (Menn and Obler, 1990; Hagiwara, 1995; Friedmann, 1998, 2001). The difficulty posed by embedded constructions is explained by tree pruning, as full relative clauses and sentential embeddings require the CP, and when the CP is unavailable, these structures are impaired. Interestingly, embedded sentences that do not require the CP, such as reduced relatives (“I saw the boy crying”), are produced correctly by individuals with agrammatism.

Tree pruning and the inaccessibility of CP cause a deficit in another important set of structures, questions. Seminal treatment studies by Shapiro, Thompson, and their group (e.g., Thompson and Shapiro, 1995) show that persons with agrammatic aphasia cannot produce well-formed Wh questions. Other studies show that in English, both Wh and yes/no questions are impaired, but in languages such as Hebrew and Arabic, Wh questions are impaired but yes/no questions are intact (Friedmann, 2002). Again, these dissociations between and within languages are a result of the unavailability of the CP: agrammatic aphasics encounter severe difficulties when trying to produce Wh questions because Wh morphemes (who, what, where, etc.) reside at CP. Yes/no questions in Hebrew and Arabic, on the other hand, do not require any overt element in CP (“You like hummus?”), and this is why they are produced correctly.

The tree pruning hypothesis is also instrumental in describing different degrees of agrammatism severity. Clinical work shows that some individuals with agrammatism use a wider range of syntactic structures than others and retain more abilities, such as verb inflection, whereas other individuals use mainly simple sentences and substitute inflections. It is possible to characterize the milder impairment as pruning at a higher site in the tree, at CP, whereas the more severe impairment results from pruning at a lower position, TP. Thus, the more mildly impaired individuals show impairment in Wh questions and embedding, but their ability to inflect verbs for both tense and agreement is relatively intact. More severely impaired individuals, who are impaired also in TP, show impaired tense inflection in addition to impairment in questions and embedding. In both degrees of severity, agreement inflection is intact. Crucially, no individual was found to exhibit a deficit in a low node without a deficit in higher nodes. In other words, there was no TP deficit without a deficit in CP, and no deficit in AgrP without a deficit in TP and CP.

Finally, tree pruning provides a principled explanation for the effect that treatment in one syntactic domain has on other domains. Thompson and Shapiro, for example, report that following question production treatment, their patients started using sentential embedding. This can be explained if treatment has enhanced the accessibility of the syntactic node that is common for the two structures, CP. Similarly, the decrease in verb omission that has been reported to accompany tense
infection improvement can be explained by enhanced accessibility to the inflectional node TP.

Current linguistic theory thus provides a useful toolbox to account for the complicated weave of spared and impaired abilities in speech production in agrammatic aphasia. The hierarchical structure of the syntactic tree enables an account of the highly selective syntactic deficit in agrammatic production in terms of syntactic tree pruning.

—Naama Friedmann

References


Further Readings


Trace Deletion Hypothesis

Some aphasic syndromes implicate grammar. Known for almost a century to be grammatically impaired in speech production (see SYNTACTIC TREE PRUNING), individuals with Broca’s and Wernicke’s aphasia are now known to have receptive grammatical deficits as well. The trace deletion hypothesis (TDH) is a collection of ideas about the proper approach to linguistic deficits subsequent to brain damage in adults, particularly Broca’s aphasia. Support for a grammatical interpretation of aphasia comes from a dense body of research from
many laboratories and varied test paradigms, administered to large groups of aphasic speakers of a variety of languages.

This article explains why a precise characterization of receptive deficits in aphasia is important, and how we came about to know it. It first aims to demonstrate that Broca’s and Wernicke’s aphasia (and perhaps other syndromes) are syntactically selective disorders in which the line dividing impaired from preserved syntax is fine and amenable to a precise characterization.

The comprehension of sentences containing grammatical transformations is impaired in aphasia. Transformations, or syntactic movement rules, are important intrasentential relations. In their various theoretical guises, these rules are designed to explain dependency relations between positions in a sentence. As a first pass, we sacrifice precision for clarity, and say that transformationally moved constituents are found in noncanonical positions (e.g., a subject in a passive sentence, The teacher was watched _______ by the student; a questioned element in object questions, Which man did Susan see _______?). Each of these sentences has a nontransformational counterpart (e.g., The student watched the teacher; Which man saw Susan?). Roughly, in transformationally derived sentences, a constituent is phonetically present in one (bolded) position but is interpreted semantically in another, “empty” position (______), which is annotated by a t (for trace of movement) and serves as a link between the semantic role of the moved constituent and its phonetic realization. Traces are therefore crucial for the interpretation of transformational sentences.

When aphasic comprehension is tested on the contrast between transformational and nontransformational sentences, a big difference is found: Broca’s aphasics understand active sentences, subject questions, and the like normally, yet fail on their transformational counterparts. This has led to the claim that in receptive language, Broca’s aphasics cannot represent traces of movement. These traces are deleted from the syntactic representations they build, hence TDH. This generalization helps localize grammatical operations in the brain. Furthermore, the highly selective character of this deficit has major theoretical ramifications for linguistic theory and the theory of sentence processing.

How exactly does the TDH work? Let us look at receptive tasks commonly used in research on brain-language relations. These are sentence-picture-matching (SPM), grammaticality judgment (GJ), and measures of reaction time (RT) during comprehension. We can consider the SPM method first. A sentence containing two arguments is presented (e.g., The student watched the teacher); it is “semantically reversible”—the lexical content does not reveal who did what to whom. The task, however, requires exactly that: subjects are usually requested to choose between two action pictures in which the roles are reversed (student watching teacher, teacher watching student). Reliance on syntax is critical. On being given a series of such sentences, Broca’s aphasics perform at above chance levels on nontransformational sentences and at chance levels on transformationally derived sentences.

What do such findings mean? Standard principles of error analysis dictate that a binary choice design allows for three logically possible outcomes: above-, below-, and at-chance performance levels. Above-chance performance is virtually normal; below-chance amounts to systematic reversal of roles in SPM; at-chance means guessing, as if the subject were tossing a coin prior to responding. Indeed, the response pattern of Broca’s aphasics on transformational sentences resembles guessing behavior. That said, the burden on the TDH is to provide a deductive explanation of why left anterior lesions, which allow normal comprehension of sentences without transformations, lead to guessing when transformational movement is at issue.

The foregoing was a simplified introductory discussion. In reality, things are somewhat more complicated. The deletion of a trace in certain cases does not hinder performance in aphasia. Our characterization needs refinement. The leading idea is to view aphasic sentence interpretation as a composite process—an interaction between incomplete syntax (i.e., representations lacking traces) and a compensatory cognitive strategy. The interpretation of moved constituents, as we saw, depends crucially on the trace; without traces, the semantic role of a moved constituent cannot be determined. Moved constituents (bolded) are thus uninterpretable. A non-linguistic, linear order–based cognitive strategy is invoked to try and salvage uninterpreted NPs (NP₁ = agent; NP₂ = theme). In English, constituents moved to a clause-initial position will thus be agents. In certain cases (e.g., subject questions), the strategy compensates correctly: in the subject relative the man who t loves Mary is tall, the head of the relative clause, the man, is moved, and receives its semantic interpretation (or thematic role) via the trace. A deleted trace blocks this process, and the strategy is invoked, assigning agent-hood to the man and yielding the correct semantics: NP₁ (the man) = agent by strategy, and NP₂ (Mary) = theme by the remaining grammar. In other cases the TDH system results in error: In object relatives—the man who Mary loves t is tall—the agent role assigned by the strategy (acting subsequent to trace deletion) gives rise to a misleading representation: NP₁ (the man) = agent by strategy, and NP₂ (Mary) = agent by the grammar. The result is a semantic representation with two potential agents for the depicted action, which leads the patients to guessing behavior. The selective nature of the aphasic comprehension deficit is captured, which is precisely what the TDH is designed to explain.

Languages with structural properties different from English lend further support to the TDH. Thus, Chinese, Japanese, German and Dutch, Spanish, and Hebrew have different properties, and the performance of Broca’s aphasics is determined by the TDH as it interacts with the particular grammar of each language. Japanese active sentences, for example, have two configurations:
[Taro-ga Hanako-o nagutta (Taro hit Hanako)—Subject Object Verb], [Hanako-o Taro-ga t nagutta—Object Subject t Verb]. These simple structures mean the same and are identical on every dimension except movement. Broca’s aphasics are above chance in comprehending the former and at chance level on the former, in keeping with the TDH. A similar finding is obtained in Hebrew. In Chinese, an otherwise SVO language like English, heads of relative clauses (annotated by the subscript h) follow the relative (unlike English, in which they precede it, as the example shows):

(1) [t zhuei gou] de mau<sub>k</sub> hen da
chase dog that cat very big
“the cat<sub>k</sub> that [t chased the dog] was very big”

(2) [mau zhuei t] de gou<sub>k</sub> hen xiao
cat chased that dog very small
“the dogs that [the cat chased t] was very small”

This structural contrast leads to remarkable contrastive performance in Broca’s aphasia: in Chinese subject relatives (1), the head of the relative (the cat) moves to the front of the sentence, lacks a role, and is assigned agent by the strategy, which leads to a correct representation in which the cat indeed chases the dog. In Chinese, the head (mau) also moves, yet to sentence-final position, and the strategy (incorrectly) assigns it the theme role. This representation now has two themes, the dog and the cat, and the result is guessing. Similar considerations that hold for the object relatives are left to the reader. English and Chinese thus yield mirror-image results, which correlate with a relevant syntactic contrast between the two languages. Further intriguing cross-linguistic contrasts exist.

Moving to other experimental tasks, we encounter remarkable cross-methodological consistency. When the detection of a violation of grammaticality critically depends on traces, the TDH predicts disability on the part of Broca’s aphasics. This is indeed the case. When given GJ tasks, they show differential performance between sentences with and without traces, again in keeping with the TDH. Finally, the rich literature on timed language reception in Broca’s aphasia (RT) suggests that on-line computation of trace-antecedent relations is compromised.

This type of deficit is not restricted to Broca’s aphasia. On at least some tests, Wernicke’s aphasics perform like Broca’s aphasics. There are contrasts between the two groups, to be sure, yet contrary to past views, there are overlapping deficits. Although this new picture is just beginning to unfold, independent evidence from functional imaging of normal populations (MRI) supports it. Both Broca’s and Wernicke’s regions of the healthy brain are involved in transformational analysis, although likely in different ways.

This series of cross-methodological findings from different languages and populations, and the TDH as a generalization over them, have a host of important theoretical implications. They show that the natural classes of structures that linguists assume have a firm neurobiological basis; they afford an unusual view on the inner workings of the human sentence processing device; they connect localized neural tissue and linguistic concepts in a more detailed way than ever before. Finally, there appears to be clinical (therapeutic) value to the view that Broca’s aphasia and Wernicke’s aphasia are grammatically selective: although currently experimental, preliminary results suggest that therapeutic methods guided by this view are somewhat efficacious.

—Yosef Grodzinsky

Further Readings


Shapiro, L. P., and Thompson, C. K. (1994). The use of linguistic theory as a framework for treatment studies in
aphasia. In P. Lemme (Ed.), Clinical aphasiology (vol. 22). Austin, TX: Pro-Ed.

Part IV: Hearing
Amplitude Compression in Hearing Aids

In the latter part of the 1980s, wide dynamic range compression (WDRC) amplification was introduced into the hearing aid market. Within a few years it was widely recognized as a fundamentally important new amplification strategy. Within 10 years nearly every hearing aid manufacturer had developed a WDRC product.

Compression is useful as a processing strategy because it compensates for the loss of cochlear outer hair cells, which compress the dynamic range of sound within the cochlea. Sensorineural hearing loss is characterized by loudness recruitment, which results from damage to the outer hair cells. WDRC compensates for this hair cell disorder, ideally restoring the limited dynamic range of the recruiting ear to that of the normal ear. This article reviews the history of loudness research, loudness recruitment, cochlear compression effects (such as the upward spread of masking) that result from and characterize OHC compression, and finally, outer hair cell physiology. The WDRC processing strategy is explained, and a short history of the development of WDRC hearing aids is provided.

Compression and Loudness

Acoustical signal intensity is defined as the flow of acoustic energy in watts per meter squared (w/m²). Loudness is the perceptual intensity, measured in either sones or loudness units (LU). One sone is defined as the loudness of a 1 kHz tone at 40 dB SPL, while 1 LU is defined as the loudness at threshold. Zero loudness corresponds to zero intensity.\(^1\)

For the case of pure tones, one sone is \(\approx 975\) LU. Isoloudness intensity contours were first determined in 1927 by Kingsbury (Kingsbury, 1927; Fletcher, 1929, p. 227). Such curves describe the relation between equally loud tones (or narrow bands of noise) at different frequencies. The intensity of an equally loud 1 kHz tone is called the loudness level, which has units of phons, measured in w/m². In 1923 Fletcher, and again in 1924 Fletcher and Steinberg, published the first key papers on the measurement of the loudness for speech signals (Fletcher, 1923a; Fletcher and Steinberg, 1924). In the 1924 paper the authors state

\[
10^{-\alpha/30} = \int_0^\infty \varphi(f) 10^{-\alpha(f)/30} df
\]

... the use of the above formula involved a summation of the cube root of the energy rather than the energy.

where \(\alpha\) is the relative intensity in dB SL, \(\varphi(f)\) is the “effective” loudness level, and \(\varphi(f)\) is an empirically determined frequency weighting factor. This cube root dependence had first been described by Fletcher the year before (Fletcher, 1923a). Fletcher and Steinberg concluded that it became apparent that the non-linear character of the ear’s transmitting mechanism was playing an important part in determining the loudness of the complex tones (p. 307).

Power law relations between the intensity of the physical stimulus and the psychophysical response are examples of Stevens’ law. Fletcher’s 1923 loudness growth equation, which for tones was found to be

\[L(I) \propto I^{1/3}\]

where \(L\) is the loudness and \(I\) is the acoustic intensity, established the important special case of Stevens’ law for sound intensity and pure-tone loudness. Their method is described in the caption of Figure 1. We now know that Fletcher and Steinberg were observing the compression induced by the cochlear outer hair cells (OHCs).

Loudness Additivity

In 1933 Fletcher and Munson published their seminal paper on loudness. This paper detailed (1) the relation of isoloudness across frequency (loudness level, or phons); (2) their loudness growth argument, described below; (3) a model showing the relation of masking to loudness; and (4) the basic idea behind the critical band (critical ratio).

Regarding the second point, rather than thinking directly in terms of loudness growth, they tried to find a formula to describe how the loudnesses of several stimuli combine. From loudness experiments with low- and high-pass speech and complex tones, and other unpublished experiments over the previous 10 years, they showed that, across critical bands, loudness (not intensity) adds. Fletcher’s working hypothesis (Fletcher and Steinberg, 1924) was that each signal is nonlinearly compressed in narrow bands (critical bands) by the cochlea, neurally coded, and the resulting band rates are added.\(^2\) The 1933 experiment clearly showed how loudness (i.e., the neural rate, according to Fletcher’s model) adds. Fletcher and Munson also determined the cochlear compression function \(G(I)\) described below for tones and speech. We now know that this function dramatically changes with sensorineural hearing loss.

Today this model concept is called loudness additivity. Their hypothesis was that when two equally loud tones are presented together but separated in frequency so that they do not mask each other, the result is “twice as loud.” The verification of this assumption lies in the predictive ability of the additivity assumption. For example, they showed that 10 tones that are all equally

\(^1\)Fletcher and Munson (1933) were able to measure the loudness below the single pure-tone threshold by using 10 equally loud tones. This proves that the loudness at threshold is not zero (Baus, Musch, and Florentine, 1998).

\(^2\)There seems to be some confusion about what is added within critical bands. Clearly, pressure must add within a critical band, or else we would not hear beats. Many books and papers assume that intensity adds within each critical band. This is true in the ensemble sense for random signals, but such a scheme will not work for tones on a single trial basis.
loud (they will be at different intensities, of course), when played together, are 10 times louder, as long as they do not mask each other. As another example, Fletcher and Munson found that loudness additivity held for signals “between the two ears” as well as for signals “in the same ear.” When the tones masked each other (namely, when their masking patterns overlapped), additivity still holds, but over an attenuated set of patterns (Fletcher and Munson, 1933). Their 1933 model is fundamental to our present understanding of auditory sound processing.

The Method. A relative scale factor (gain) $a$ may be defined either in terms of the pressure or in terms of the intensity. Since it is the voltage on the earphone that is scaled, the most convenient definition of $a$ is in terms of the pressure, $P$. It is typically expressed in dB, given by $20 \log_{10}(a)$.

Two equally loud tones were matched in loudness by a single tone scaled by $a^*$. The asterisk indicates this special value of $a$. The resulting definition of $a^*$ is given by

$$L(a^*P) = 2L(P),\quad (1)$$

which says that, when the single tone pressure, $P$, is scaled by $a = a^*$, the loudness, $L(a^*P)$, is twice as loud as the unscaled signal. Given the relative loudness level (in phons) of “twice as loud,” defined by $a^*(I)$, the loudness growth function $G(I)$ may be found by graphical methods or by numerical recursion, as shown in Fletcher (1953, p. 190) and in Allen (1996b). The values of $a^*(I)$ found by Fletcher in different papers published between 1933 and 1953 are shown in Figure 2.

The Result. These two-tone loudness matching experiments showed that for $f_1$ between 0.8 and 8.0 kHz, and $f_2$ far enough away from $f_1$ (above or below) so that 17% of its original energy. The corresponding relative level for 1 kHz high-pass-filtered speech is 7%. These functions are shown as the solid lines in the figure. The high- and low-pass loudnesses do not add to 1 since the two solid lines cross at about 11%. After taking the cube root, however, the loudness curves cross at 50% (i.e., at 0.8 kHz, $0.125^{1/3} = 0.5$), and therefore sum to 100%. A level of 11.3 $\mu$ BARS (dynes/cm²) corresponds to 1.13 Pa, which is close to 95 dB SPL. (From Fletcher, 1929, p. 236.)

![Figure 1. Effect of low- and high-pass filtering on the speech loudness level. The broadband speech is varied in level until it is equal in loudness to the low-pass-filtered speech. This is repeated for each value of the filter cutoff frequency. The experiment was then repeated for the high-pass speech. The percent reduction of the equally loud broadband speech energy is plotted against the filter cutoff frequency. For example, if broadband speech is to be equal in loudness to speech that has been low-pass-filtered to 1 kHz, it must be reduced in level to

![Figure 2. Loudness and $a^*$ as a function of the loudness level, in phons. When $a^*$ is 9 dB, loudness increases as the cube root of intensity. When $a^*$ is 3 dB, loudness is proportional to intensity. (From Fletcher, 1953.)](image-url)
there is no masking, the relative level \( x \) was found to be \( \approx 9 \, \text{dB} \) (ca. 1953) for \( P_1 \) above 40 dB SPL. This value decreased linearly to 2 dB for \( P_1 \) at 0 phons, as shown in Figure 2.

From this formulation, Fletcher and Munson found that at 1 kHz, and above 40 dB SPL, the pure-tone loudness \( G \) is proportional to the cube root of the signal intensity \( [G(I) = (P/P_{ref})^{2/3}] \) because \( x^* = 2^{3/2} \) (9 dB). Below 40 dB SPL, loudness was frequently assumed to be proportional to the intensity \( [G(I) = (P/P_{ref})^2, x^* = 2^{1/2}, or 3 \, \text{dB}] \). Figure 2 shows the loudness growth curve and \( x^* \) given in Fletcher (1953, p. 192, Table 31) as well as the 1938 and 1933 papers. As may be seen from the figure, in 1933 they found values of \( x \) as high as 11 dB near 55 dB SL. Furthermore, the value of \( x^* \) at low levels is not 3 dB but is closer to 2 dB. Fletcher’s statement that loudness is proportional to intensity \( (x^* \text{ is } 3 \, \text{dB near threshold}) \) was an idealization that was appealing, but not supported by actual results.

**Recruitment and the Rate of Loudness Growth**

Once loudness had been quantified and modeled in 1933 by Fletcher and Munson, Mark Gardner, a close personal friend and colleague of Harvey Fletcher, began measuring the loudness growth of hearing-impaired subjects. In about 1934 Gardner first discovered the effect that has become known as loudness recruitment (Gardner, 1994), first reported by Steinberg and Gardner (1937) as well as the 1938 and 1933 papers. As may be seen from the figure, in 1933 they found values of \( x \) as high as 11 dB near 55 dB SL. Furthermore, the value of \( x^* \) at low levels is not 3 dB but is closer to 2 dB. Fletcher’s statement that loudness is proportional to intensity \( (x^* \text{ is } 3 \, \text{dB near threshold}) \) was an idealization that was appealing, but not supported by actual results.

**Loudness Growth in the Recruiting Ear.** Figure 3 shows a normal loudness growth function along with a simulated recruiting loudness growth function. It is necessary to plot these functions on a log-log (log loudness versus dB SPL) scale because of the dynamic ranges of loudness and intensity. The use of the dB and log loudness has resulted in a misinterpretation of the rate (slope) of recruitment. In the figure we see that for a 4 dB change in intensity about 58 dB SPL, the loudness changes by 1.0 sone in the normal ear and 0.5 sones in the recruiting ear. While the slope looks steeper on a log plot, the actual rate of loudness growth (in sones) in the recruiting ear is smaller. Its misdefinition as “the abnormally rapid growth of loudness” has lead to some serious conceptual errors about loudness and hearing loss. Correct statements about loudness recruitment include “the abnormal growth of loudness” or “the abnormally rapid growth of relative loudness \( \Delta L/L \) (or log loudness).”

**Fowler’s Mistake.** After learning from Wegel about the yet unpublished recruitment measurements of Steinberg and Gardner, E. P. Fowler attempted to use recruitment to diagnose middle ear disease (Fowler, 1936). In cases of hearing loss involving financial compensation, Fowler stated that recruitment was an “ameliorating” factor (Fowler, 1942). In other words, he viewed recruitment as a recovery from hearing loss—its presence indicated a reduced hearing loss at high intensities. Thus, given two people with equal threshold losses, the person having the least amount of recruitment was given greater financial compensation (the loss could be due to middle ear disease, and the individual would receive greater compensation than someone having a similar sensorineural loss).

In my view, it was Fowler’s poor understanding of recruitment that led to such terms as complete recruitment versus partial recruitment and hyper-recruitment. Complete recruitment means that the recruiting ear and the normal ear perceive the same loudness at high intensities. Steinberg and Gardner described such a loss as a variable loss (i.e., sensorineural loss) and partial recruitment as a mixed loss (i.e., having a conductive component that acts as a frequency-dependent fixed
attenuation). They, and Fowler, verified the conductive component by estimating the air-bone gap.

Steinberg and Gardner attempted to set the record straight. They clearly understood what they were dealing with, as is indicated in the following quote (Steinberg and Gardner, 1937, p. 20):

Owing to the expanding action of this type of loss it would be necessary to introduce a corresponding compression in the amplifier in order to produce the same amplification at all levels.

This model of hearing and hearing loss, along with the loudness models of Fletcher and Munson (1933), are basic to the eventual quantitative understanding of cochlear signal processing and the cochlea’s role in detection, masking, and loudness in normal and impaired ears. The work by Fletcher (1950) and Steinberg and Gardner (1937), and work on modeling hearing loss and recruitment by Allen (1991) support this view.

Compression and Masking

In 1922, one year after publishing the first threshold measurements with Wegel, Fletcher published measurements on the threshold of hearing in the presence of a masking tone (Fletcher, 1923a, 1923b). Wegel and Lane’s classic and widely referenced paper on masking, and their theory of the cochlea, soon followed, in 1924. In Figure 4 we reproduce one of the figures from Fletcher’s 1923 publication (which later appeared in the 1924 Wegel paper) showing the upward spread of masking due to a 400 Hz tone. As we shall see, these curves characterize the nonlinear compressive effects of outer hair cell compression.

Critical Band Masking. When the probe is near the masker in frequency, as in the case of the 0.45 kHz probe tone shown in Figure 4, the growth of masking is close to linear. Such near-linear growth is called Weber’s law. The masked-threshold probe intensity \( I_p \) is equal to the masker intensity \( I_m \) plus 1 JND, namely

\[
I_p(I_m) = I_m + \Delta I(I_m).
\]

The masking appears to be linear because the relative JND (e.g., \( \Delta I/I \approx 0.1 \)) is small. As the intensity of the masker is increased, the variations in the JND \( \Delta I(I_m) \) with respect to the masker intensity \( I_m \) appear negligible, making \( I_p(I_m) \) appear linear. Weber’s law is therefore observed when the probe is within a critical bandwidth of the masker. One sees deviations from Weber’s law when plotting more sensitive measures, such as \( \Delta I(I_m)/I_m \) (Riesz, 1928).

Upward Spread of Masking. The suppression threshold, \( I_p^*(I_m, f_p) \), is defined as the smallest masker intensity such that the slope of \( I_p^*(I_m, f_p) \) with respect to \( I_m \) is greater than 1. Since the probe slope is close to 2.4 dB/ dB over a range of intensities, this threshold is best estimated from the intercept of the \( I_p^*(I_m, f_p) \) regression line with the abscissa. For the 3 kHz probe, the suppression threshold intensity is 60 dB SL. Such suppression is only seen for probes greater than the masker frequency (\( f_p > f_m \)). For probes that are sufficiently higher in frequency than the masker (e.g., \( f_p \geq 2 \) kHz in Fig. 4), the masking is close to zero dB SL until the masker intensity reaches the suppression threshold at about 50–60 dB SL. In other words, the masked threshold, defined as the intensity where the masking of the probe begins, and the suppression threshold are nearly the same. The suppression threshold for the dashed-line, superimposed on the “solid-fat” \( f_p = 3 \) kHz probe curve in Figure 4, is 60 dB SL; its slope is 2.4 dB/dB. For every 1 dB increase in

\[\text{An asterisk is used to indicate that the intensity is at threshold.}\]
the masker intensity $I_m$, the probe threshold intensity $I_p^s(I_m, f_m, f_p)$ must be increased by 2.4 dB to return it to its detection threshold (Delgutte, 1990; Allen, 1997c). Namely, above $I_p^s(f_p = 3$ kHz, $f_m = 0.4$ kHz) = 60 dB SL (i.e., $I_m > 10^6I_m$),

$$\frac{I_p^s(I_m)}{I_m^{2.4}} = \left(\frac{I_m}{I_p^s}\right)^{2.4}.$$  \hspace{1cm} (2)

From Figure 4, a surprising and interesting crossover occurs near 65–70 dB for the 1 kHz probe. As highlighted by the dashed box, the 1 kHz probe threshold curve crosses the 0.45 kHz probe threshold curve. At high levels, there is more masking at 1 kHz than at the probe frequency. This means that the masker excitation pattern peak has shifted toward the base of the cochlea (i.e., toward the stapes). Follow-up forward masking studies have confirmed this observation (Monson and Gardner, 1950, Fig. 8). McFadden (1986) presents an excellent and detailed discussion of this interesting “half-octave shift” effect that is recommended reading for all serious students of hearing loss.

**Downward Spread of Masking.** For probes lower than the masker frequency (Fig. 4, 0.25 kHz), while the threshold is low, the masking is weak, since it has a slope that is less than linear. This may be explained by the migration of the more intense high-frequency (basal) masker excitation pattern away from the weaker probe excitation pattern (Allen, 1999b).

**The Physiology of Compression**

What is the source of Fletcher’s tonal cube root loudness growth (i.e., Stevens’ Law)? Today we know that the basilar membrane motion is nonlinear in intensity, as first described by Rhode in 1971, and that cochlear OHCs are the source of the basilar membrane nonlinearity. The history of this insight is both interesting and important.

In 1937, Lorente de Nó theorized that abnormal loudness growth associated with hearing loss (i.e., recruitment) is due to hair cell damage (Lorente de Nó, 1937). From noise trauma experiments on humans one may conclude that recruitment occurs in the cochlea (Carver, 1978). Animal experiments have confirmed this prediction and have emphasized the importance of OHC loss (Liberman and Kiang, 1978; Liberman and Dodds, 1984). This loss of OHCs causes a loss of the basilar membrane compression (Pickles, 1982, p. 287). It follows that the cube root tonal loudness growth starts with the nonlinear compression of basilar membrane motion due to stimulus-dependent voltage changes within the OHC.

**Two-Tone Suppression.** The neural correlate of the 2.4 dB/dB psychoacoustic suppression effect (the upward spread of masking) is called two-tone suppression (2TS) (Sachs and Abbas, 1974; Fahey and Allen, 1985; Delgutte, 1990). Intense low-frequency tones attenuate low-level high-frequency tones to levels well below their threshold. The close relationship between the two effects has only recently been appreciated (Allen, 1997c, 1999b). The 2TS and upward spread of masking (USM) effects are important to the hearing aid industry because they quantify the normal cochlear compression that results from OHC processing. To fully appreciate the USM and 2TS, we need to describe the role of the OHC in nonlinear cochlear processing. In Figure 5 the operation of USM/2TS is summarized in terms of neural excitation patterns.

**Cochlear Nonlinearity: How?**

We still do not know precisely what controls the basilar membrane nonlinearity, although we know that it results from OHC stiffness and length changes (He and Dallos, 2000), which are a function of the OHC membrane voltage (Santos-Sacchi and Dilger, 1987). This voltage is determined by shearing displacement of the hair cell cilia by the tectorial membrane (TM). The most likely cause of nonlinear basilar membrane mechanics is changes in the micromechanical impedances within the organ of
Corti. This conclusion follows from ear canal impedance measurements, expressed in terms of nonlinear power reflectance, defined as the retrograde to incident power ratio (Allen et al., 1995). In a transmission line, the reflectance of energy is determined by the ratio of the load impedance at a given point divided by the local characteristic impedance of the line. It is this ratio that is level dependent (i.e., nonlinear).

Two Models. It is still not clear how the cochlear gain is reduced, and that is the subject of intense research. There are two basic but speculative theories. The first is a popular but qualitative theory, referred to as the cochlear amplifier (Kim et al., 1980). The second is a more physical and quantitative theory that requires two basic assumptions. The first assumption is that the tectorial membrane acts as a bandpass filter on the basilar membrane signal (Allen, 1980). The second assumption is that the OHCs dynamically “tune” the basilar membrane (i.e., the cochlear partition) by changing its net stiffness, causing a dynamic migration in the characteristic place with intensity (Allen, 1997b). Migration is known to occur (McFadden, 1986), so this assumption is founded on experimental dogma.

We cannot yet decide which, if either, of these two theories is correct, but for the present discussion, it is not important. The gain of the inner hair cell (IHC) cilia excitation function is signal dependent, compressing the 120 dB dynamic range of the acoustic stimulus to less than 60 dB. When the OHC voltage becomes depolarized, the OHC compliance increases, and the characteristic frequency (CF) of the basilar membrane shifts toward the base, reducing the nonlinear wide dynamic range compression.

Cochlear Nonlinearity: Why?

The discussion above leaves unanswered why the OHCs compress the signal on the basilar membrane. The answer to this question has to do with the large dynamic range of the ear. In 1922 Fletcher and Wegel were the first to use electronic instruments to measure the threshold and upper limit of human hearing (Fletcher and Wegel, 1922a, 1922b), thereby establishing the 120 dB dynamic range of the cochlea.

The IHCs are the cells that process the sound before it is passed to the auditory nerve. Based on the Johnson (thermal) noise within the IHC, it is possible to accurately estimate a lower bound on the RMS voltage within the IHC. From the voltage drop across the cilia, we may estimate the upper dynamic range of the cell. The total dynamic range of the IHC must be less than this ratio, or less than 65 dB (e.g., 55–60 dB) (Allen, 1997b). The dynamic range of hearing is about 120 dB. Thus, the IHC does not have a large enough dynamic range to code the dynamic range of the input signal. Spread-of-excitation models and neuron threshold distribution of neural rate do not address this fundamental problem. Nature’s solution to this problem is the OHC-controlled basilar membrane compression.

The formula for the Johnson RMS thermal electrical membrane noise voltage $|V_c|$ due to cell membrane leakage currents is given by $3 \langle |V_c|^2 \rangle = 4kTB$, where $B$ is the cell membrane electrical bandwidth, $k$ is Boltzmann’s constant, $T$ is the temperature in degrees Kelvin, and $R$ is the cell membrane leakage resistance. The cell bandwidth is limited by the membrane capacitance $C$. The relation between the cell RC time constant, $\tau = RC$, and the cell bandwidth is given by $B = 1/\tau$, leading to

$$|V_c| = \sqrt{\frac{4kT}{C}}. \tag{3}$$

The cell membrane capacitance $C$ has been determined to be about 9.6 $pF$ for the IHC (Kros and Crawford, 1990) and 20 $pF$ for the OHC. From Equation 3, $V_c = 21$ $\mu V$ for IHCs at body temperature ($T = 310^\circ K$).

Although the maximum DC voltage across the cilia is 120 mV, the maximum RMS change in cell voltage that has been observed is about 30 mV (I. J. Russell, personal communication). The ratio of 30 mV to the noise floor voltage (21 $\mu V$), expressed in dB, is 63 dB. Thus it is impossible for the IHC to code the 120 dB dynamic range of the acoustic signal. Because it is experimentally observed that, taken as a group, IHCs do not code a wide dynamic range, the nonlinear motion of the basilar membrane must be providing compression within the mechanics of the cochlea prior to IHC detection (Allen and Neely, 1992; Allen, 1997a).

Summary. Based on a host of data, the physical source of cochlear hearing loss and recruitment is now clear. The dynamic range of IHCs is limited to about 50 dB. The dynamic range of the sound level at the eardrum, however, is closer to 100–120 dB. Thus, there is a difficulty in matching the dynamic range at the drum to that of the IHC. This is the job of the OHCs.

It is known that OHCs act as nonlinear elements. For example, the OHC soma axial stiffness, $K_{ohc}$, depends directly on the voltage drop across the cell membrane, $V_{ohc}$. As the OHC cilia excitation is varied from “soft” to “loud,” the OHC membrane voltage is depolarized, causing the cell to increase its compliance (and length). The result is compression due to a decrease in the IHC (cochlear) signal gain.

Multiband Compression

During the two decades from 1965 to 1985, the clinical audiological community was attempting to answer the question: Are compression hearing aids better than a well-fitted linear hearing aid? A number of researchers concluded that linear fitting is always superior to compression. When properly adjusted, linear filtering is close to optimum for speech whose level has been adjusted for

$^3$While the thermal noise is typically dominated by the shot noise, the shot noise is more difficult to estimate. Since we are trying to bound the dynamic range, the thermal noise is a better choice for this purpose. The shot noise reduces the dynamic range further, strengthening the argument.
optimum listening. Papers that fall in this category include Braida et al. (1979) and Lippmann et al. (1981). However, Lippmann et al. are careful to point out the flaw in preadjusting the level (see p. 553).

Further criticisms were made by Plomp (1988, 1994), who argued that compression would reduce the modulation depth of the speech. However, compression of a broadband signal does not reduce the modulations in sub-bands.

All these results placed the advocates of compression in a defensive minority position. Villchur vigorously responded to the challenge of Plomp, saying that Plomp’s argument was flawed (Villchur, 1989). The filter bandwidths used in WDRC hearing aids are not narrow enough to reduce the modulations in critical bandwidths. Other important papers arguing for compression include Hickson (1994), Killion (1996a, 1996b), Killion et al. (1990), and Mueller and Killion (1996). A physiology paper that is frequently cited in the compression literature is Ruggero and Rich (1991).

Other work that found negative results used compression parameters that were not reasonable and time constants that were too slow. Long time constants with compression produce very different results and are not in the category of syllabic compression. Such systems typically have artifacts, such as noise “pumping,” or they simply do not react quickly enough to follow a lively conversation. Imagine, for example, a listening situation with a quiet and a loud talker having a conversation. In this situation, the compressor gain must operate at syllabic rates to be effective. The use of multiple bands ensures that a signal in one frequency band does not control the gain in another band. Slow-acting compression (AGC) may be fine for watching television, but not for conversational speech. Such systems might be viewed as a replacement for a volume control (Dillon, 1996, 2001; Moore et al., 1985; Moore, 1987).

A key advocate of compression was Ed Villchur, who critically recognized the importance of Steinberg and Gardner’s observations on recruitment as a loss of compression. He vigorously promoted the idea of compression amplification hearing aids. Personally supporting the cost of the research with dollars from his very successful loudspeaker business, he contracted David Blackmer of dbx to produce a multiband compression hearing aid for experimental purposes. Using his experimental multiband compression hearing aid, Villchur experimented on hearing-impaired individuals, and found that Steinberg and Gardner’s observations and predictions were correct (Villchur, 1973, 1974). Villchur clearly articulated the point that a well-fitted compression hearing aid improved the dynamic range of audibility and that what counted, in the end, was audibility. In other words, “If you can’t hear it, you can’t understand it.” This had a certain logical appeal.

Fred Waldhauer, a Bell Labs analog circuit designer of some considerable ability, heard Villchur speak about his experiments on multiband compression. After the breakup of the Bell System in 1983, Waldhauer proposed to AT&T management that Bell Labs design and build a multiband compression hearing aid as an internally funded venture. Eventually Bell Labs built a digital wearable hearing aid prototype. It quickly became apparent that the best processing strategy compromise was a two-band compression design that was generally similar to the Villchur scheme. With my colleague Vincent Pluvinage, we designed digital hardware wearable hearing aids, and with the help of Joe Hall and David Berkley of AT&T, and Patricia Jeng, Harry Levitt, Arlene Newman, and many others from City University of New York, we developed a fitting procedure and ran several field trials (Allen et al., 1990). AT&T licensed its hearing aid technology to ReSound on February 27, 1987.

Unlike today, in 1990 multiband compression was widely unaccepted, both clinically and academically (Dillon, 2001). Why is this? It was, and remains, difficult to show quantitatively the nature of the improvement of WDRC. It is probably fair to say that only with the success of ReSound’s WDRC hearing aid in the marketplace has the clinical community come to accept Villchur’s claims. It may be possible to clarify the acceptance issue by presenting two common views of what WDRC is and why it works. One’s adopted view strongly influences how he or she thinks about compression. They are the articulation index (AI) view and the loudness view.

The articulation index view is based on the observation that speech has a dynamic range of about 30 dB in one-third octave frequency bands (French and Steinberg, 1947). The assumption is that the AI will increase in a recruiting ear as the compression is increased, if the speech is held at a fixed loudness. This view has led to unending comparisons between the optimum linear hearing aid and the optimum compression hearing aid.

The loudness view is based on restoring the natural dynamic range of all sounds to provide the impaired listener with all the speech cues in a more natural way. Soft sounds for normals should be soft for the impaired ear, and loud sounds should be loud. According to this view, loudness is used as an index of audibility, and complex arguments about JNDs, speech discrimination, and modulation transfer functions just confound the issue. This view is supported by the theory that OHCs compress the IHC signals.

Neither of these arguments deals with important and complex issues such as changing of the critical band with hearing loss, or the temporal dynamics of the compression system. Analysis of these important details is interesting only after the signals are placed in the audible range.

Summary

This article has reviewed the early research on loudness, loudness recruitment, and masking, which are relevant to compression hearing aid development. The outer hair cell is damaged in sensorineural hearing loss, and this causes the cochlea to have reduced dynamic range.

When properly designed and fitted, WDRC has proved to be the most effective speech processing
strategy we can presently provide for sensorineural hearing loss compensation. It works because it supplements the OHC compressors, which are damaged with sensorineural hearing loss.

Acknowledgments

I would especially like to thank one anonymous review, Harvey Dillon, Brent Edwards, Ray Kent, Mead Killon, Harry Levitt, Ryuji Suzuki, and Ed Villchur.

—Jont B. Allen

References


Liberman, M., and Dodds, L. (1984). Single neuron labeling and chronic cochlear pathology: III. Stereocilia damage and


### Assessment of and Intervention with Children Who Are Deaf or Hard of Hearing

The purpose of communication assessment of children with educationally significant hearing loss differs from the purpose of assessing children with language or learning disabilities. Since the diagnosis of a hearing disability has already been made, the primary goal of communication assessment is to determine the impact of the hearing loss on language, speech, auditory skills, or cognitive, social-emotional, educational and vocational development, not to diagnose a disability. It is critical to determine the rate of language and communication development and to identify strategies that will be most beneficial for optimal development.

Plateaus in language development at the 9–10-year age level, in reading development at the middle third grade to fourth grade level (Holt, 1993), and in speech intelligibility at about 10 years (Jensema, Karchmer, and Trybus, 1978) have been reported in the literature. The language plateaus appear to be the result of developmental growth, which ranges from 43%–53% for children with profound hearing loss using hearing aids (Boothroyd, Geers, and Moog, 1991; Geers and Moog, 1988) to 60%–65% of the normal range of development for children with severe loss using hearing aids (Boothroyd, Geers, and Moog, 1991) and for children with profound hearing loss using cochlear implants (Blamey et al., 2001). In contrast, in a study of 150 children, Yoshinaga-Itano et al. (1998) reported that children with hearing loss only who were early-identified (within the first 6 months of life) had mean language levels at 90% of the rate of normal language development through the first 3 years of life. A study of children in Nebraska (Moeller, 2000) reported similar levels of language development (low-average range) for a sample of 5-year-old children receiving early intervention services in the first 11 months of life. Later-identified children were able to achieve language development commensurate with the early-identified/intervened group when their families were rated as “high parent involvement” in the intervention services.

With the advent of universal newborn hearing screening, the population of children who are deaf or hard of hearing will change rapidly during the next decade. By 2001, 35 states had passed legislation to...
establish universal newborn hearing screening programs, five additional states had legislation in progress, and five states had established programs without legislation. The age at which hearing loss is identified should drop dramatically throughout the United States, and this drop should be accompanied by intervention services, beginning in the newborn period. In the state of Colorado, infants referred from UNHS programs are being identified with hearing loss at 6–8 weeks of age and enter into intervention programs almost immediately thereafter.

For the population of children identified with hearing loss within the first 2 months of life, baseline communication assessments are typically conducted at 6-month intervals during the first 3 years of life (Stredler-Brown and Yoshinaga-Itano, 1994; Yoshinaga-Itano, 1994). Almost all of the infant assessment instruments are parent questionnaires that address the development of receptive and expressive language (e.g., MacArthur Communicative Development Inventories, Minnesota Child Development Inventory, Vineland Social Maturity Scales), auditory skills, early vocalizations, cognitive, fine motor, gross motor, self-help, and personal-social/social-emotional issues. Videotaped analysis of parent-child interaction style and spontaneous speech and language production is frequently included.

Spontaneous speech samples should be analyzed to identify the number of different consonant phones. The number of different consonant phones produced in a spontaneous 30-minute language sample taken in the home between 9 months and 50 months of age is a good predictor of speech intelligibility (Yoshinaga-Itano and Sedey, 2000). The primary development of speech for children with mild to moderate hearing loss is concentrated between 2 and 3 years of age, while the preschool years, ages 3–5 years, are a significant growth period for children with moderate to severe hearing loss. Speech development for children with profound hearing loss who use hearing aids is very slow in the first 5 years of life. Although 75% of children with mild through severe hearing loss achieved intelligible speech by 5 years of age, only 20% of children with profound hearing loss who used conventional amplification achieved this level by age 5. Level of expressive language and degree of hearing loss were the two primary predictors of speech intelligibility.

Videotaped interactions of parent-child communication can be analyzed for maternal bonding and emotional availability (Pressman et al., 1999), turn-taking (Musselman and Churchill, 1991), use of pause time, maintenance of topic, topic initiation, attention-getting devices, the development of symbolic play, symbolic gesture, and communication intention strategies (comments, requests, answers, commands) of both the parent and the child (Yoshinaga-Itano, 1994). These analyses provide important information for the family and intervention provider to help design strategies for optimal development. Reciprocal relationships have been reported. Parents adjust language as their child’s language improves (Cross, Johnson-Morris, and Nienhuys, 1985), and low levels of maternal turn control are associated with greater gains in expressive language (Musselman and Churchill, 1991). At present, studies of causality have been insufficient to determine the efficacy or superiority of various intervention strategies. However, some interventions that are theoretically grounded and are characteristic of programs that demonstrate optimal language development are parent-centered, provide objective developmental data, assist parental decisions about methodology, based on the developmental progress of the individual child, include a strong counseling component aimed at reducing parental stress and assisting parents in the resolution of their grief, and provide guidance in parent-child interaction strategies.

The language development of early-identified (within the first 6 months of life) children with hearing loss is similar to their nonverbal development, particularly in regard to symbolic play development (Snyder and Yoshinaga-Itano, 1999; Yoshinaga-Itano and Snyder, 1999; Mayne et al., 2000). About 60% of the variance in early language development is predicted by nonverbal cognitive measures such as symbolic play and age at identification of hearing loss. Mode of communication, degree of hearing loss, socioeconomic status, ethnicity, and sex were not shown to predict language development. These results contrast sharply with the school-age literature, in which race, ethnicity, and socioeconomic status are primary predictors of reading achievement (Holt, 1993).

In order to maintain the successful language development of the early-identified children, the purpose of evaluation in the first 5 years of life should be to monitor and chart the longitudinal developmental progress of the child, with two primary goals: (1) to achieve language development commensurate with nonverbal cognitive development, and (2) to achieve language development in children with hearing loss only, within the normal range of development. In an analysis of almost 250 children, an 80% probability of language within the low-normal range in the first 5 years of life was reported if a child identified with hearing loss had been born in a Colorado hospital with a universal newborn hearing screening program (Yoshinaga-Itano, Coulter, and Thomson, 2000).

In addition to an analysis of communication skills, cognitive development, and age at identification, assessments should include information about the social-emotional development of both parents and children. Relationships have been found between language development and parental stress (Pipp-Siegel, Sedey, and Yoshinaga-Itano, 2002), emotional availability (Pressman et al., 1999) parent involvement (Moeller, 2000), grief resolution (Pipp-Siegel, 2000), development of sense of self (Pressman, 2000), and mastery motivation (Pipp-Siegel et al., 2002). The relationships examined in these studies do not establish causes, but it is plausible that intervention strategies focused on these areas may enhance the language development of young children with hearing loss. Counseling strategies with parent sign language instruction enhanced language development.
among children using total communication (Greenberg, Calderon, and Kusche, 1984).

Assessment strategies during the preschool period, ages 3–5 years, continue to focus on receptive and expressive vocabulary, but their primary focus shifts to the development of syntax and morphology and pragmatic language skills (Yoshinaga-Itano, 1999). Spontaneous language sample analysis is recommended for expressive syntax analysis. There is a shift from parent questionnaire developmental assessments and assessments of spontaneous communication to clinical or school-based elicited assessments at this age period. Testers need to ensure full access to the information being presented, either through fully functioning amplification devices, adequate speech reading accessibility, or the skills of a fluent sign.

During the school-age period, standardized assessments consist of regularly administered tests of language (receptive and expressive vocabulary, syntax, pragmatics, and phonology), reading and writing, mathematics, other content areas, and social-emotional development (Yoshinaga-Itano, 1997). Researchers hypothesize that there may be as many as four possible routes to literacy for children with hearing loss: (1) spoken language to printed language decoded to speech, (2) English-based signs to printed English, (3) American Sign Language (ASL) to print, with English-based signs as an intermediary, and (4) ASL to print (Musselman, 2000). Assessments of knowledge of English semantics, syntax, and phonological processing, accuracy and speed of word identification, and orthographic encoding should be included, since several studies have found that these variables are significantly related to reading comprehension (Musselman, 2000). Among children who use sign language, finger-spelling ability and general language competence in either ASL or English should be included (Musselman, 2000). For all children, assessments should also focus on the metalinguistic and metacognitive strategies (knowing how to use and think about language) used by the students in person-to-person and written communication (Gray and Hosie, 1996). “Theory of mind” assessments provide information about the cognitive ability of the student to understand a variety of different perspectives (Strassman, 1997). Students need to develop strategies to acquire and elaborate world knowledge, elaborate vocabulary knowledge (both conversational and written), and to use this knowledge to make inferences in social, communicative interactions and reading/academic situations (Paul, 1996; Jackson, Paul, and Smith, 1997).

—Christine Yoshinaga-Itano

References


Research on auditory development intensified in the mid-1970s, expanding our understanding of auditory processes in infants and children. Recent comprehensive reviews of this literature include those by Aslin, Jusczyk, and Pisoni (1998), Werner and Marean (1996), and Werner and Rubel (1992). Two recurrent issues are whether the limitations of testing methods cause true abilities to be underestimated, and whether auditory development is substantially complete during infancy or continues into childhood.

**Absolute Sensitivity**

Auditory sensitivity is within about 10–15 dB of adult values by 6 months after birth, but behavioral and electrophysiological methods show different trends during early infancy. Behavioral thresholds at 1 month are approximately 40 dB above adult values, with substantial improvement through 6 months and gradual improvement through the early school years (Schneider et al., 1986; Olsho et al., 1988). By 6 months the audibility curve (threshold plotted against frequency) is adultlike in shape, although sensitivity continues to improve into childhood. Several behavioral tests are used to assess infant hearing, including conditioned head turning by infants older than about 6 months and observation of subtle behavioral responses to sounds in younger infants. It is of clinical significance that for infants younger than 4–6 months, behavioral methods lack reliability for individual assessment (Bourland, Tharpe, and Ashmead, 2000). Infant-adult differences reflect true sensory differences as well as nonsensory factors such as attention, but at least in older infants only about 4 dB of the infant-adult difference is attributable to nonsensory factors (Nozza and Henson, 1999). Electrophysiological measures (primarily the auditory brainstem response, ABR) reflect nearly adultlike hearing sensitivity from early infancy (Werner, Folsom, and Mancel, 1994). Although electrophysiological measures are less susceptible than behavioral methods to inattention and off-task behavior, another reason for the discrepancy between methods is that the ABR reflects only peripheral processing. That is, in early infancy, auditory sensitivity may be constrained by neural immaturity central to the processes measured by the ABR.

**Frequency Resolution**

Like absolute sensitivity, frequency resolution approaches adult values by 6 months after birth. The segregation of auditory processing by frequency is demonstrated by the enhanced effectiveness of masking sounds that are closer in frequency to the signal. Infants’ cochlear frequency resolution has been studied via suppression tuning curves from otoacoustic emissions. Such frequency resolution is adultlike at full-term birth (Bargones and Burns, 1988; Abdala, 1998). Beyond the cochlea, resolution for higher frequencies (ca. 4 kHz) is not finely tuned for the first 3 months after birth, but by 6 months resolution is adultlike, as seen in behavioral and ABR masking studies (Spetner and Olsho, 1990; Abdala and Folsom, 1995). Some studies indicate improvement in frequency resolution up to age 4 years, but Hall and Grose (1991) showed that when certain methodological issues are taken into account, children have adultlike frequency resolution by 4 years. Despite the early development of frequency resolution, selective attention based on frequency is not well developed even by 9 months (Bargones and Werner, 1994).

**Temporal Processing**

There may be a more protracted course for temporal processing than for some other aspects of auditory development. Adults detect acoustic gaps of 2–5 ms, whereas infants require gaps up to ten times longer (Werner et al., 1992). However, infant gap detection may be adult-like when tested in conditions that minimize adaptation effects (Trehub, Schneider, and Henderson, 1995; Trainor et al., 2001). Gap detection improves through 5–10 years (Wightman et al., 1989). Hall and Grose (1994) found that older children and adults had
similar time constants for detection of amplitude modulation, but that preschool-aged children needed larger modulations. Discrimination between small differences in duration improves during infancy and into middle childhood (Jensen and Neff, 1993). The coarseness of temporal acuity in infants and children is surprising, since typically developing children do well at linguistic processing. But individual differences in infants’ temporal resolution are associated with language development at 2 years (Trehub and Henderson, 1996). Besides temporal acuity, another consideration is how sound is integrated across time, reflected by improvement of detection thresholds as a sound remains on longer. Adult detection improves as the signal extends up to ¼ s, whereas infants have steeper improvement curves extending to longer durations (Berg and Boswell, 1995). A related finding is that forward masking operates over longer masker-signal intervals in 3-month-olds than in 6-month-olds or adults, implying that a temporal aspect of masking is mature by about 6 months (Werner, 1999).

**Intensity Resolution and Loudness**

Early studies of intensity discrimination showed that neonates and young infants detect large intensity changes. Sinnott and Aslin (1985) reported intensity difference limens for 7- to 9-month-olds of 3–12 dB, compared to 1.8 dB for adults. Other studies indicate that between birth and 12 months, the intensity difference limens ranges from 2 to 5 dB (Bull, Eilers, and Oller, 1984; Tarquinio, Zelazo, and Weiss, 1990). Although infants are good at discriminating changes in intensity, this ability may not be adultlike until 6 years of age (Jensen and Neff, 1993). Research on loudness perception by infants has not been reported, but there has been some work with children. Studies using magnitude estimation and cross-modality matching have demonstrated that 4- to 7-year-olds described the growth of loudness of a tone similarly to adults (Collins and Gescheider, 1989). Serpanos and Gravel (2000), also using cross-modality matching paradigms, reported that loudness growth functions of children 4–12 years of age were similar to those of adults.

**Pitch and Music Perception**

Research on pitch and music perception suggests that by late in the first year of life, infants extract relational information across frequencies. At 7 months, infants categorize octave tonal complexes by fundamental frequency, even if the fundamental is missing and cochlear distortions are ruled out (Clarkson and Clifton, 1995; Montgomery and Clarkson, 1997). A substantial literature on music perception in infants shows sensitivity to melodic contours, even when transposed across octaves (Trehub, 2001). Although findings on pitch and music suggest integrative processing, other work in nonmusical contexts indicates that infants are better able to discriminate absolute than relative spectral differences (Saffran and Griepentrog, 2001).

**Speech Perception**

As early as the 1960s, investigators proposed that speech was processed by a separate perceptual module of a larger language system unique to humans. Although speech is still considered a special signal, the existence of an innate and specialized speech-processing mode unique to humans remains in question (Aslin, Jusczyk, and Pisoni, 1998). Since the initial studies of infant speech perception (Eimas et al., 1971), many investigators have shown that mechanisms used to perceive speech are in place long before infants utter their first words at around 12 months of age. For example, infants can discriminate between most English speech sound pairs much as adults do (Kuhl, 1987), based on parameters such as voice onset time, manner, and place of articulation. Despite infants’ ability to categorize speech sounds, the boundaries between categories continue to sharpen into early childhood (Nittouer and Studdert-Kennedy, 1987). Also, children age 3–4 years are not as adept as older children and adults at perceiving speech sounds from brief-onset information (Ohde and Haley, 1997). Infants also can generalize phoneme recognition across different talkers, fundamental frequencies, and positions within a syllable (Eilers, Wilson, and Moore, 1977; Miller, Younger, and Morse, 1982). The ability to categorize vowels, also known as equivalence classes, has been demonstrated in infants as young as 2 months of age (Marean, Werner, and Kuhl, 1992). Changes in speech perception as a result of postnatal experience occur within the first year of life, as shown by selective perception of speech sounds from the native language (see Werner and Marean, 1996, chap. 6). As early perception of linguistic input is so remarkable, any obstacle to the receipt of audible and clear speech sounds by an infant, such as hearing loss, should be suspected of adversely affecting language development.

**Spatial Hearing**

Sound localization is rather crude in early infancy, improving nearly to adult levels by 1 year and slowly thereafter. Newborns turn their heads toward sound sources, but this response wanes, reappearing in brisk form around 4 months (Clifton, 1992). This trend, along with the onset of response to the precedence effect (related to reflected sounds) at 4 months, suggests an increasing role of the auditory cortex in sound localization. The low responsivity to sounds from birth to 4 months makes it difficult to assess hearing ability behaviorally in this age range, as noted above in the discussion of auditory sensitivity. Precision of sound localization, as measured by the minimum audible angle, improves from 20°–25° at 4 months to 8°–10° at 12 months, with further gradual changes to adultlike values of 1°–3° by 5 years (Ashmead, Clifton, and Perris, 1987; Morrongiello, Fenwick, and Chance, 1990). This improvement may entail integration across several sound localization cues, since sensitivity to single cues is better than predicted from localization (Ashmead et al., 1991). A phenomenon related to spatial hearing is maskin
level differences, which are smaller in children less than 7 years old than in older children or adults (Grose, Hall, and Dev, 1997). Regarding distance, 6-month-olds distinguish between sounds within reach versus those beyond reach, even if sound pressure level is removed as a distance cue (Litovsky and Clifton, 1992).

See also physiological bases of hearing; pediatric audiology: the test battery approach; hearing loss screening: the school-age child.

—Daniel H. Ashmead and Anne Marie Tharpe

References


The success of that early attempt led to the development of a more sophisticated multichannel ABI device (Brackmann et al., 1993; Shannon et al., 1993; Otto et al., 1998, 2002). The first commercial multichannel ABI was developed in a collaborative effort between the House Ear Institute, Cochlear Corporation, and the Huntington Medical Research Institutes. The multichannel ABI was approved by the U.S. Food and Drug Administration in October 2000.

Several commercial ABI devices are available, all basically similar to the original device. ABIs are virtually identical in design to cochlear implants except for the electrode assembly (Fig. 1). The electrode assembly is a flat, paddle-like structure with platinum electrical contacts along one side. The overall size of the assembly is generally 2–3 mm × 8 mm and is designed to fit within the lateral recess of the IV ventricle. The electrical contacts are 0.5–1.0 mm in diameter, which is sufficient for keeping the electrical charge density at the stimulated neurons within safe limits (Shannon, 1992). All ABI devices have an external speech processor unit that contains a microphone to pick up the acoustic sound, a signal processor to convert the acoustic sound to electrical signals, and a transmitter/receiver to send the signals across the skin to the implanted portion of the device. The implanted unit decodes the received signal and produces controlled electrical stimulation of the electrodes. The most widely used ABI electrode array (Fig. 2; manufactured by Cochlear Corp.) consists of 21 platinum disk contacts, each 700 µm in diameter. The contacts are placed in three rows along a Silastic rubber carrier that is 8 mm × 2.5 mm (Fig. 3).

Present ABI electrodes are designed to be placed within the lateral recess of the IV ventricle. Anatomical studies (Moore and Osen, 1979) of the human cochlear nucleus complex and imaging studies of early ABI patients demonstrated that this location produced the most effective auditory results and the fewest nonauditory side effects (Shannon et al., 1997; Otto et al., 1998, 2002). Electrical stimulation in the human brainstem can potentially produce activation of many nonauditory structures (cranial nerves VII, IX, X, and cerebellum, for

**Auditory Brainstem Implant**

People who are deafened by bilateral acoustic tumors are not candidates for a cochlear implant because tumor removal often severs the auditory nerve. The auditory brainstem implant (ABI) is designed for those patients. It is intended to bypass the auditory nerve and electrically stimulate the human cochlear nucleus.

Neurofibromatosis type 2 (NF2) is a genetic disorder that causes multiple tumors of the cranial nerves and spinal cord, among other symptoms. The gene causing NF2 has been located on chromosome 22 (Rouleau et al., 1993; Trofatter et al., 1993). The defining symptom of the disease is bilateral tumors originating on the Schwann cells of the vestibular branch of nerve VIII (vestibular schwannomas). These tumors are life-threatening, and their removal usually produces bilateral deafness. Patients with NF2 cannot benefit from a cochlear implant because they have no auditory nerve that can be stimulated from an intracochlear electrode. In 1979 William House and William Hitzeberger attempted to provide auditory sensations for an NF2 patient by placing a single pair of electrodes in the cochlear nucleus following tumor removal (Edgerton, House, and Hitzelberger, 1984; Eisenberg, et al., 1987).
example). Fortunately, the human cochlear nucleus complex almost completely surrounds the opening of the lateral recess of the IV ventricle, and the levels of current delivered to the ABI are not large enough to activate other brainstem nuclei more than about 2 mm away (Shannon, 1989, 1992; Shannon et al., 1997).

Vestibular schwannomas can be visualized and removed via several surgical approaches, of which the retrosigmoid and translabyrinthine approaches are the most common. The translabyrinthine approach allows better visualization of the mouth of the lateral recess following tumor removal and thus better access for placement of the ABI electrode array (Brackmann et al., 1993).

Of the first 80 patients implanted with the multichannel ABI at the House Ear Institute, 86% received sufficient auditory sensations that they could use the ABI in daily life. For most ABI users the primary benefit is as an aid to lipreading, since only a few ABI patients can understand words with the ABI without lipreading. For most patients the present ABI device functions at a level similar to that of single-channel cochlear implants, even when many electrodes are used in the ABI speech processor. ABI patients are able to detect sound and are able to discriminate sounds based on coarse temporal properties (Shannon and Otto, 1990; Otto et al., 2002). On average, ABI patients receive a 30% improvement in speech understanding when the sound from the ABI is added to lipreading alone (Fig. 4). A few ABI patients (10%) achieve significant word and sentence recognition with the device, and a few (4/80) can actually converse in a limited fashion on the telephone.

Most ABI patients perceive variations in amplitude and temporal cues but receive little, if any, spectral cues. ABI patients are able to perceive changes in pitch with changes in pulse rate only up to about 150 Hz, which is about an octave lower than observed for cochlear implant listeners and for temporal pitch discrimination for normal-hearing listeners. Typically, the dynamic range of amplitude for the ABI is only 6 dB or less in terms of electrical current range. We estimate that ABI patients may be able to discriminate only 10 amplitude steps within their dynamic range, in contrast to 20–40 steps for cochlear implants and 200–250 steps for acoustic hearing. Some ABI patients have relatively small differences in pitch across their electrode array, while others show a large change in pitch. In general, patients who do better at speech recognition tend to have a larger pitch range across their electrode array, but not all patients who have a large pitch range have good speech recognition.

Most ABI patients have some electrodes that cause nonauditory side effects. Almost all of these nonauditory effects are benign and produce tingling sensations along the body on the side ipsilateral to the ABI. Nonauditory sensations are produced from stimulation of the cerebellar flocculus (which causes a sensation of eye jitter) and from antidromic activation of the cerebellar peduncle. In patients with an intact facial nerve on the implanted side there is a chance of activation of the facial nerve, causing facial tingling and even motor activation. In most cases, electrodes that produce nonauditory side effects are simply turned off and not stimulated.

One of the possible reasons for the limited success of the ABI is that the present electrode is placed on the surface of the cochlear nucleus. Unfortunately, the tonotopic axis of the human cochlear nucleus is orthogonal to the surface of the nucleus (Moore and Osen, 1979), and thus orthogonal to the axis of the electrode array as well. To obtain better access to the tonotopic dimension of the human cochlear nucleus requires penetrating electrodes (McCreery et al., 1998). A penetrating electrode ABI system is presently under development. Its initial trial is anticipated for 2002.

—Robert V. Shannon
References


Auditory Brainstem Response in Adults

The auditory brainstem response (ABR) is a series of five to seven neurogenic potentials, or waves, that occur within the first 10 ms following acoustic stimulation (Sohmer and Feinmesser, 1967; Jewett, Romano, and Williston, 1970). The potentials are the scalp-recorded synchronous electrical activity from groups of neurons in response to a rapid-onset (<1 ms) stimulus. An example of these potentials, with their most common labeling scheme using Roman numerals, is shown in Figure 1. Waves I and II are generated in the auditory nerve, wave III is predominantly from the cochlear nucleus, and waves IV and V are predominantly from the superior olivary complex and lateral lemniscus (Moller, 1993).
The ABR is valuable in audiology and neurology because of its reliability and highly predictable changes in many pathological conditions affecting the auditory system. The ABR may be used in a number of ways in adults. Perhaps the most common use is in the diagnostic assessment of a hearing loss, either to determine the site of the lesion or to determine the function of the neural system. A second use is to estimate auditory sensitivity in patients who are unable or unwilling to provide accurate behavioral thresholds. A third use is for monitoring the auditory nerve and brainstem pathways during surgery for auditory nerve tumors or vestibular nerve section. For diagnostic and neural monitoring purposes, the latencies and amplitudes of the most reliable waves, waves I, III, and V, are analyzed. For estimation of auditory sensitivity, the lowest detection level for wave V is used to approximate auditory threshold. Responses are replicated to ensure reliability of the waveforms.

Because the potentials are small (<1 μV) and embedded in high levels of background electrical noise, several techniques are used to enhance the visibility of the potentials (ASHA, 1988). Surface electrodes are attached to the scalp of the patient, with the placement in line with the orientation of the dipoles of the neural generators. Pairs of electrodes are used such that one electrode picks up positive activity and one electrode picks up negative activity from the neural generators. Differential amplification/common mode rejection enhances the electrical activity that is different between two electrodes (the potentials) and reduces the activity that is the same between electrodes (the random background electrical noise). The physiological response is filtered to eliminate the extraneous electroencephalographic activity and external line noise. Signal averaging increases the magnitude of the time-locked response and minimizes the random background activity. Artifact rejection eliminates unusually high levels of electrical activity that are impossible to eliminate through averaging.

The diagnostic interpretation of the ABR is based on the latencies and amplitudes of the component waves. The latencies, or times of occurrence, of the waves are the more reliable measures because latencies for a given person remain stable across recording sessions unless intervening pathology has occurred. The latency is evaluated in absolute and relative terms. Absolute latency is measured from the arrival of the stimulus at the ear to the positive peak of waves I, III, and V. Relative latency is measured between relevant peaks within the same ear or between ears. The absolute latency reflects the state of the auditory system to the generation site of each wave but may be affected by conductive or cochlear pathology, making it difficult to isolate any delay from neural pathology. Interpeak intervals allow an estimation of neural conduction time and are less dependent on peripheral pathology than absolute latencies are. Interaural values limit variability by using the nonsuspect ear as a control in a patient with unilateral pathology, but interaural values also may be affected by asymmetrical conductive or cochlear pathology.

The amplitudes of the waves are more variable than the latencies and therefore are less useful for determining normality of the waveform. Amplitudes, measured from the positive peak to the averaged baseline or from positive peak to subsequent negative trough, are affected by the quality of the electrode contact, physiological noise levels of the patient, and amplitudes of adjacent waves. Consequently, absolute amplitude values are little used except in the case of absent waves. Relative amplitudes, particularly the amplitude ratio of waves I and V, may be useful measures to control for measurement variables. A decrement in wave V amplitude relative to wave I amplitude may suggest auditory nerve or low brainstem pathology.

For the estimation of auditory threshold, wave V may be traced to its detection threshold, which is typically defined as the lowest stimulus level at which wave V can be seen. The ABR does not measure hearing per se, but correlates with auditory sensitivity in most cases. Clicks, which are broadband stimuli, provide estimates of average sensitivity in the range of 2000–4000 Hz because of cochlear physiology that biases the response to the high-frequency region of the cochlea (Fowler and Durrant, 1993). For better definition of thresholds across the frequency range, frequency-specific stimuli, such as tone pips or clicks with ipsilateral masking, may be used (Stapells, Picton, and Durieux-Smith, 1993). Because of limitations in the signals and the cochlea, thresholds for frequencies below 1000 Hz are difficult to obtain, and other methods, such as middle or late auditory-evoked potentials or steady-state evoked potentials, may yield better results or additional information (see Hall, 1992; Jacobson, 1993).

Clicks at high presentation levels are the most commonly used stimuli for diagnostic ABRs because their abrupt rise times elicit the necessary degree of synchrony to obtain the full complement of waves. Although the normal ABR to a click is dominated by neurons associated with 2000–4000 Hz, a peripheral hearing loss at those frequencies may effectively filter the stimulus, causing different frequency regions to dominate the response in different hearing losses, which will affect the ensuing ABR latencies (see Fowler and Durrant, 1993). ABR latencies from suspect ears are compared with the norms at equivalent sound pressure levels at the cochlea.

Figure 1. The normal ABR elicited by a high-level click stimulus, with waves I–VI labeled.
The typical effects of auditory pathology on the latencies and thresholds of the ABR are discussed below.

**Normal Hearing.** Normative latencies may vary somewhat among clinics, but are typically derived from the mean and ±2 standard deviations of the latencies of waves from a jury of listeners with normal hearing. Examples of normative values for waves I, III, and V are shown as the bounded areas in the lower panel in Figure 2. Typical normal interpeak intervals are <2.51 ms for interpeak interval I–III, <2.31 ms for III–V, <4.54 ms for I–V, and <0.4 ms for interaural V latency (Bauch and Olsen, 1990). Figure 2 includes the ABRs for high-level signals to threshold for a person with normal hearing, along with the latencies of those waves plotted against the normal range. All latencies are within the normal limits, and the threshold is 40 dB peak sound pressure level (pSPL), which equals 10 dB nHL (normalized hearing level).

**Conductive Hearing Loss.** The absolute latencies are prolonged and the amplitudes are reduced relative to the degree of the conductive component of a hearing loss, but the interwave intervals are normal because the neural system is intact. Consequently, a person with a 30-dB conductive hearing loss will produce an ABR to a 90-dB nHL click that is approximately equivalent to the ABR to a 60-dB nHL click for a person with normal hearing. An example of a waveform and the resulting latencies from a 30-dB flat, conductive hearing loss are shown in Figure 3. All absolute latencies are prolonged, but the I–V latency difference is within normal limits. The threshold is 70 dB pSPL.
Cochlear Hearing Loss. The degree and configuration of the hearing loss affect the latencies of the waves in a cochlear hearing loss, although typically the I–V interval is normal because the neural system is intact. Most mild to moderate cochlear losses do not affect the latencies of the ABR for high-level click stimuli, although the amplitudes of the waves may be reduced. Severe high-frequency hearing losses may reduce the amplitudes and prolong the absolute latencies of the waves with little effect on the I–V latency interval, although wave I may be absent. An example of the waveform and resulting latencies from a moderate, flat hearing loss are shown in Figure 4. Absolute and interwave latencies are within normal limits, and the threshold is 100 dB pSPL.

Retrocochlear Hearing Loss. Retrocochlear pathology refers to any neural pathology of the auditory system that is beyond the cochlea and may include such disorders as acoustic neuromas, multiple sclerosis, brainstem strokes, and head trauma. Retrocochlear pathology may produce a variety of effects on the latencies and morphology of the ABR depending on the type, location, and size of the pathology. Effects may include absence of waves, prolonged absolute latencies or interwave intervals, or prolonged interaural wave V latencies. In Figure 5, the waveform and resulting latencies are shown from a patient with an acoustic neuroma and a mild, high-frequency hearing loss. Wave I is within normal limits, whereas wave V is significantly prolonged, yielding a prolonged interwave I–V interval. The threshold for wave V does not correspond to the behavioral auditory threshold.

Figure 4. The top panel illustrates the ABR for a person with a moderate, flat cochlear hearing loss, with wave V labeled to its visual detection threshold of 100 dB pSPL. The lower panel shows the latencies for waves I, III, and V plotted within the normative values, indicated by the bounded areas.

Figure 5. The top panel illustrates the ABR for a person with a vestibular schwannoma, with the reliable waves labeled. The lower panel shows the latencies for waves I and V plotted against the normative values for those waves. Wave I is within normal limits, whereas wave V is significantly prolonged, yielding a prolonged interwave I–V interval. The threshold for wave V does not correspond to the behavioral auditory threshold.
and latencies of waves I and V are monitored during the surgery when compromise of the nerve is a possibility. Wave I provides an index of cochlear function and neural function peripheral to the tumor and wave V provides an index of neural function central to the tumor. Significant prolongations of the latency of wave V or a decline in the amplitude of wave V may suggest damage to the nerve, which then may be at risk for permanent injury. Occasionally, the surgical technique may be modified in an attempt to avoid permanent injury to the nerve.

—Cynthia G. Fowler

References


Stapells, D., Picton, T. W., and Durieux-Smith, A. (1993). The short latency auditory brainstem response (ABR), beginning with wave I. The presence of hearing dysfunction in quiet, and of poor or absent ABR, distinguish this disorder from central

Further Readings


Auditory Neuropathy in Children

The disorder now known as auditory neuropathy (AN) has been defined only within the past 10 years (Starr et al., 1991), although references to patients with this disorder appeared as early as the 1970s and 1980s (Friedman, Schulman, and Weiss, 1975; Ishii and Toriyama, 1977; Worthington and Peters, 1980; Kraus et al., 1984; Jacobson, Means, and Dhib-Jalbut, 1986). This disorder is particularly deleterious when it occurs in childhood because it causes significant disturbance of encoding of temporal features of sound, which severely limits speech perception and, consequently, the development of oral language skills.

Patients with AN have three key characteristics. First, they have a hearing disorder in the form of elevated pure-tone thresholds (which can vary from slight to profound) or significant dysfunction of hearing in noise. Second, they have evidence of good outer hair cell function, in the form of either present otoacoustic emissions or an easily recognized cochlear microphonic component. Third, they have evidence of neural dysfunction at the level of the primary auditory nerve. This condition manifests with an abnormal or absent auditory brainstem response (ABR), beginning with wave I. The presence of hearing dysfunction in quiet, and of poor or absent ABR, distinguish this disorder from central
auditory dysfunction, in which hearing and ABR are both normal.

The presence of normal outer hair cell function and the absence of wave I of the ABR narrow the potential sites of lesion in AN to (1) the inner hair cell, (2) the synaptic junction between the inner hair cell and the auditory nerve, and (3) the peripheral portion of the auditory nerve. There is evidence to support the first and third possibilities, but no direct evidence of synaptic disorder.

Starr (2001) has found that approximately one-third of all patients with AN have symptoms of peripheral nerve disease. Approximately 80% of adults with AN demonstrate concomitant peripheral neuropathy, while no patients less than 5 years old show clinical evidence of peripheral nerve disorder. Peripheral neuropathy that was not evident in some of the younger patients emerged in children who were followed over time. In patients with other peripheral nerve involvement, disease of the primary portion of the auditory nerve would be the most parsimonious explanation for the auditory disorder.

More direct evidence of primary auditory nerve disease in humans was reported by Spoendlin (1974) from postmortem temporal bone histologic studies of two sibling adult patients with moderate hearing loss. These individuals had a full complement of inner and outer hair cells but significant loss of spiral ganglion cells. Similar findings were reported by Nadol (2001) in a patient with Charcot-Marie-Tooth syndrome and hearing loss.

Harrison (1999, 2001) has shown that both carboplatin treatment and anoxia can induce isolated inner hair lesions in chinchillas. The same animals had otoacoustic emissions and abnormal ABR results. Amatuzzi et al. (2001) recently reported an autopsy analysis of the temporal bones of three premature infants; findings included isolated inner hair cell loss with a full complement of outer hair cells and auditory neurons. These infants had failed ABR screening while in the neonatal intensive care unit. This study presented the first evidence in humans that an isolated inner hair cell disorder is a possible explanation for AN.

No currently available clinical tools can provide data to distinguish between the inner hair cell and the auditory nerve as site of lesion in AN. Because young children with AN often do not show evidence of other peripheral nerve involvement, it is particularly difficult to know what the underlying pathology of their AN might be. However, it is also not clear how any distinction in pathology could be used to remediate the hearing difficulties in these patients.

Description of Patients

Most patients with AN have disease onset in childhood (Sininger and Oba, 2001). The sex distribution is approximately equal. Close to half of patients with AN have a family history or an associated genetic syndrome, indicating a genetic basis for the disorder. In some cases, specific chromosomal anomalies have been isolated (Butinar et al., 1999; Kalaydjieva et al., 1996, 1998). Other syndromes that have been associated with AN include Charcot-Marie-Tooth syndrome, Friedreich's ataxia, Ehlers-Danlos syndrome, and Stevens-Johnson syndrome (Sininger and Oba, 2001).

Many patients with AN show no risk factors. However, other health issues often associated with AN in infants include hyperbilirubinemia and prematurity (Stein et al., 1996; Sininger and Oba, 2001). As we begin to detect hearing loss in the newborn period with screening, it will be possible to obtain more data on those factors that may be directly associated with AN.

The pure-tone hearing loss in patients with AN ranges from slight to profound, with a greater percentage of severe and profound loss than in patients with sensory hearing loss. Any configuration of loss can occur, but low-frequency emphasis or flat configurations are most often seen. Only rarely is AN unilateral, but such cases have been described.

The symptoms and hearing loss of patients with AN can fluctuate dramatically on a day-to-day basis or more frequently. The most dramatic instance of rapid changes in symptoms has been described as “temperature-sensitive auditory neuropathy.” Starr et al. (1998) described three such patients in whom severe symptoms, including nearly complete loss of hearing and ABR, accompanied a slight fever. The symptoms in these patients would remit, leaving nearly normal auditory function, as soon as the core temperature returned to normal.

Symptoms of AN can progress slowly over time, and the symptoms seen in newborns sometimes improve in the first few months of life.

Patients with AN have poor performance on all auditory tasks involving temporal processing. For example, AN patients have dramatically abnormal gap detection thresholds and temporal modulation transfer functions when compared with patients with sensory hearing loss or normal hearing. In contrast, loudness functions, such as intensity discrimination and temporal integration, are normal for subjects with AN (Zeng et al., 2001).

Speech perception ability is notably reduced in these patients, especially in regard to the degree of hearing loss (Sininger and Oba, 2001). Poor speech discrimination can be directly linked to reduced temporal processing (Zeng et al., 2001).

The ABR is abnormal or absent in cases of auditory neuropathy. In all cases except profound loss, the threshold of the ABR does not correspond to the auditory thresholds, and for this reason the ABR cannot be used to predict hearing levels in children with AN. In some cases of AN, a wave V can be distinguished, but it is usually small in amplitude, extended in latency, and the response threshold is unrelated to pure-tone hearing levels.

The ABR waveform of the patient with AN, when recorded with electrodes placed at the mastoid or otherwise near the ear, will usually show evidence of a large cochlear microphonic (CM) component. This pattern can be easily distinguished from an early neural response because the CM waveform will invert with stimulus
Figure 1. Audiologic findings in a 7-year-old boy with auditory neuropathy. A, Audiogram. Speech awareness threshold: right ear = 20 dB HL, left ear = 25 dB HL. Speech discrimination: right ear = 28%, left ear = 8%. Tymanometry was within normal limits; the acoustic reflex threshold was absent. B, Transient evoked otoacoustic emission. C, Auditory brainstem response. Study was performed using insert earphones, with a click stimulus rate of 25/s at 80 dB nHL. Rarefaction and condensation graphs are overlaid; right ear and left ear.
polarity and demonstrate a mirror image when the waveforms corresponding to the two stimuli are superimposed. Also, unlike neural response, the CM component does not change latency with a decrease in stimulus level and is relatively undisturbed by noise masking. The CM can be evidence of either inner or outer hair cell function. The CM can be distinguished from stimulus artifact by a simple procedure. If an insert earphone is used, clamping the tubing will cause the CM to disappear, because no effective stimulus will reach the ear. At the same time, stimulus artifact from the transducer will remain after the tubing is clamped.

In most cases of AN, an otoacoustic emission (OAE) is seen, regardless of the degree of hearing loss. However, the OAE has diminished over time in some of these patients, for unknown reasons (Deltenre et al., 1997). The CM component can be substituted for the OAE as evidence of hair cell function.

Brainstem reflexes involving the auditory system, including middle ear muscle reflexes and olivocochlear reflexes (suppression of OAE with contralateral noise), are almost universally absent in patients with AN. Again, unlike in sensory loss, this finding is true regardless of the degree of hearing loss present.

Rehabilitation Strategies

Inherently poor timing in the auditory system of patients with AN requires specialized approaches to the development of speech and language. The use of visual information to supplement speech perception is most important. Manual communication in a total communication system (oral speech and sign language simultaneously) is often recommended. For very young children with no language system, this approach helps to ensure that a language system will develop regardless of the capacity of the auditory system to process oral speech. In mild cases of AN (those with mild or moderate hearing thresholds), a cued-speech approach can be useful (Cone-Wesson, Rance, and Siningger, 2001).

Standard amplification does not provide the same degree of benefit for patients with AN as in patients with conductive or sensory hearing loss. However, for many children, some advantages can be gained from amplification, including lower thresholds for the detection of environmental sounds and even small increases in speech perception ability. Parents should be cautioned that the benefits of amplification will be limited, and a trial period should be used to determine if any help is afforded by the hearing aids.

Cochlear implants have been used in many children with AN (Trautwein, Siningger, and Nelson, 2000; Shallop et al., 2001; Trautwein et al., 2001). Electrical stimulation of the auditory nerve may reintroduce the temporal encoding through neural synchrony, necessary for speech perception. Most of the children with AN who have received cochlear implants perform similarly to other deaf children, but good performance is not always achieved (Trautwein et al., 2001). An important question is whether children with moderate hearing loss and AN who do not receive benefit from hearing aids should be considered for cochlear implants. Results in patients with moderate loss due to AN and cochlear implants may be available in the near future to help answer this important question.

Case Report

In a 7-year-old boy with AN, the medical history was unremarkable. Speech development was normal before the age of 5 years, but after that time, response to sound was inconsistent, and difficulty with speech production was noted. At age 7 years the audiogram in Figure 1 was obtained, revealing a moderate hearing loss bilaterally. Tympanometry was normal, acoustic reflexes were absent, and otoacoustic emissions were present. ABR testing revealed no response to 80 dB nHL stimuli, but evidence of a cochlear microphonic component was seen in the recording. Magnetic resonance imaging of the brain and cranial nerves VII and VIII was normal, as was a neurological examination. Amplification and FM systems were used sparingly, with little success. This child relies heavily on speechreading, supplemented with manual communication as necessary.

This child represents a possible late-onset case of AN of unknown etiology. His audiogram shows the often seen nonuniform configuration with peaks and valleys. Significant fluctuations in his hearing have been noted over time. He is also typical in that he relies very heavily on visual cues, including speechreading, supplemented by signs, for receptive communication.

—Yvonne S. Siningger

References


Auditory Scene Analysis

Sensory systems such as hearing probably evolved in order for organisms to determine objects in their environment, allowing them to navigate, feed, mate, and communicate. Objects vibrate and as a result produce sound. An auditory system provides the neural architecture for the organism to process sound, and thereby to learn something about these objects or sound sources. In most situations, many sound sources are present at the same time, but the sound from these various sources arrives at the organism as one complex sound field, not as separate, individual sounds. The challenge for the auditory system is to process this complex sound field so that the individual sound sources can be determined. That is, the auditory system is presented with a complex auditory scene, and the auditory images in this scene are the sound sources (Bregman, 1990). The auditory system must be capable of performing auditory scene analysis if it is to determine the sources of sounds.

Auditory scene analysis is not undertaken in the auditory periphery (the cochlea and auditory nerve). The auditory periphery provides a spectral-temporal neural code for the acoustic information contained within the auditory scene. That is, the auditory nerve relays to the auditory brainstem the coding performed by the cochlea. This neural code provides the central nervous system with information about the spectral components that make up the auditory scene in terms of their frequencies, levels, and timing. The central auditory nervous system must then analyze this peripheral neural code so that the individual sound sources that generated the scene can be determined.

What information might be preserved in the peripheral code that is usable by the central auditory system for auditory scene analysis? Several cues have been suggested. They include frequency separation, temporal separation, spatial separation, level differences in spectral profiles, asynchronous onsets and offsets, harmonic structure, and temporal modulation (Yost, 1992). These are properties of the sound generated by sound sources that may be preserved in the peripheral code. As an example, if two sound sources each vibrate with a different frequency, the two frequencies will be mixed into a single auditory scene arriving at the listener. The auditory system could ascertain that two frequencies are present, indicating two sound sources. We know that this is possible since within certain boundary conditions, the
auditory periphery codes for the frequency content of any complex sound.

The example of two sound sources each producing a different frequency is the basis of a set of experiments designed to investigate auditory scene analysis. Imagine that the sound coming from each of the two sources is pulsed on and off so that the sound from one source (one frequency) is on when the sound from the other source (a different frequency) is off. The perception of the stimulus condition could be described as a single sound with an alternating pitch. However, since each sound could be from a different source, the perception of the stimulus condition could also be that of two sound sources, each producing a pulsing tone of a particular frequency. Each of these perceptions is possible given the exact frequencies and timing used in the experiment. When the perception is one of two different sound sources, the percept is often described as if two perceptual streams were running side by side, each stream representing a sound source. The stimulus conditions that lead to this form of stream segregation are likely to be those that promote the segregation of sound from one source from that of another source (Bregman, 1990). Many of the parameters listed above have been studied using this auditory streaming paradigm. In general, stimulus parameters associated with frequency are more likely to support stream segregation (Kubovy, 1987), but most of the parameters listed can support auditory stream segregation under certain conditions.

Experiments to study auditory scene analysis, such as auditory streaming, require listeners to process sound over a large range of frequencies and over time. Since a great deal of work in auditory perception and psychoacoustics has concentrated on short-time processing in narrow frequency regions, less is known about auditory processing across wide regions of the spectra and longer periods of time. Therefore, obtaining a better understanding of cross-spectral processing and long-time processing is very important for revealing processes and mechanisms that may assist auditory scene analysis (Yost, 1992).

One of the traditional examples of cross-spectral processing that relates to auditory scene analysis is the processing of the pitch of complex sounds, such as the pitch of the missing fundamental (see Pitch Perception). For these spectrally complex stimuli, usually ones that have a harmonic structure, a major perceptual aspect of the stimulus is the perception of a single pitch (Moore, 1997). Conditions such as the pitch of the missing fundamental suggest that the auditory system uses a wide range of frequencies to determine the complex pitch and that this complex pitch may be the defining acoustic characteristic of a harmonic sound source, such as the musical note of a piano key. A sound consisting of the frequencies 300, 400, 500, 600, and 700 Hz would most likely have a perceived pitch of 100 Hz (the missing fundamental in the sound’s spectrum). The 100-Hz pitch may help in determining the existence of a sound source with this 100-Hz harmonic structure.

The example of the missing fundamental pitch can be generalized to describe acoustic situations in which the auditory system segregates sounds into more than one source. A naturally occurring sound source is unlikely to have all but one of its frequency components harmonically related. Thus, in the example cited above, it is unlikely that a single sound source would have a spectrum with frequency components at 300, 400, 550, 600, and 700 Hz (the 550-Hz component is the inharmonic component that replaced the 500-Hz component). In fact, when one of the harmonics of a series of harmonically related frequency components is “mistuned” from the harmonic relationship (550 Hz in the example), listeners are likely to perceive two pitches (as if there were two sound sources), one associated with the 100-Hz harmonic relationship and the other with the frequency of the mistuned harmonic (Hartmann, McAdams, and Smith, 1990). That is, the 550-Hz mistuned harmonic is perceptually segregated as a separate pitch from the 100-Hz complex pitch associated with the rest of the harmonically related components. Such dual pitch perception suggests there were two potential sound sources. In this case, the auditory system appears to be using a wide frequency range (300–700 Hz) to process these two pitches, and hence perceives two potential sound sources.

The complex pitch example can also be used to address the role of stimulus onset as another potential cue used for auditory scene analysis. Two sound sources may each produce a harmonically related spectrum, such that one sound source may have frequency components at 150, 300, 450, 600, and 750 Hz (harmonics of 150 Hz) and another at 233, 466, 699, and 832 Hz (harmonics of 233 Hz). When presented in isolation these two sounds will produce pitches of 150 and 233 Hz. However, if the two stimuli are added together so that they come on and go off together, it is unlikely that two pitches will be perceived. The perception of two pitches can be recovered if one of the complex sounds comes on slightly before the other one, even though both sounds remain on together thereafter. Thus, if the sound from one source comes on (or in some cases goes off) before another sound, then the asynchronous onsets (or offsets) promote sound source segregation, aiding in auditory scene analysis (Darwin, 1981).

If two sounds have different temporal patterns of modulation they may be perceptually segregated on the basis of temporal modulation (especially if the modulation is amplitude modulated; Moore and Alcantara, 1996). Detecting a tonal signal in a wideband-noise background is improved if the noise is amplitude modulated, suggesting that the modulation helps segregate the tone from the noise background (Hall, Haggard, and Fernandes, 1984).

Sounds from spatially separated sources may help in determining the auditory scene. The ability of the auditory system to use sound to locate objects in the real world also appears to help segregate one sound source from another (Yost, 1997). However, the ability to determine the instruments (sound sources) in an orchestral
piece played over a single loudspeaker suggests that having actual sources in our immediate environment at different locations is not required for auditory scene analysis.

Thus, in order to use sound to determine something about objects in our world, the central auditory system must process the neural code for sound in order to parse that code into subsets of neural information where each subset may be the neural counterpart of a sound source. Several different parameters of sound sources are preserved by the neural code and may form the basis of auditory scene analysis. Determining the sources of sound requires processing across a wide range of frequencies and over time, and is required for organisms to successfully cope with their environments.

—William A. Yost

References


Further Readings


Auditory Training

Auditory training includes a collection of activities, the goal of which is to change auditory function, auditory behaviors, or the ways in which individuals approach auditory tasks. Auditory training most commonly is associated with the rehabilitation of individuals with hearing loss, but it has been used with other populations that have presumed difficulties with auditory processing, such as children with specific language impairment, phonologic disorder, dyslexia, and autism (Wharry, Kirkpatrick, and Stokes, 1987; Bettison, 1996; Merzenich et al., 1996; Habib et al., 1999). Auditory training has been applied to children diagnosed with central auditory processing and to adults learning a second language (Solma and Adepoju, 1995; Musiek, 1999). It also has been used experimentally to assess the plasticity of speech perceptual categories and to determine the neuroanatomical substrates of speech perception learning and organization (Werker and Tees, 1984; Bradlow et al., 1997; Tremblay et al., 1997, 1998, 2001; Wang et al., 1999).

Most auditory training programs are organized around three parameters: auditory processing approach, auditory skill, and stimulus difficulty level (Erber and Hirsh, 1978; Erber, 1982; Tye-Murray, 1998). When implementing an auditory training program, the first decision is whether to use a top-down (synthetic) or a bottom-up (analytic) processing approach or a combination of both. Most normal-hearing adults tend to rely primarily on top-down processing when listening to ongoing speech, but if the goal for a patient is to learn or relearn an auditory skill, a more bottom-up processing approach may be warranted. The skill to be learned or relearned (e.g., detection, discrimination, identification, and comprehension) and the complexity of the stimuli are largely dictated by the status of the patient’s auditory skills and the goals of auditory training. The stimulus difficulty can be manipulated by changing such factors as set size, acoustic similarity, speed, linguistic complexity, lexical familiarity, visual cues, contextual support, and environmental acoustics. It also can be adjusted by digitally manipulating specific acoustic parameters such as formant transition duration, f1 intensity, or noise burst spectra. Typically, the training starts with skill and stimulus levels at which the patient just exhibits difficulty. Then the skill and stimulus difficulty are systematically increased as performance improves until the training goal is attained. For example, the final goal might be that the patient will reduce fricative place confusions to 25% in connected discourse. In a bottom-up approach, the patient might work on fricative place discrimination in CV syllables with the same vowel, then with different vowels, words, phrasal and sentence structures, and finally at the discourse level. Speaking rate, vocal intensity, environmental acoustics, and contextual cues can be adjusted at each level to increase the listening demands. At the discourse level, the conversational
demands also can be increased systematically. As with most skills, learning likely is facilitated by cycling across a number of difficulty levels and stimuli, having frequent training sessions, and varying the duration and location of the training.

In a more top-down approach, the patient might start at the discourse level and work on evaluating and using context to predict topic and word choices. The focus is less on hearing the place cues and more on increasing the awareness of various contextual cues that can be used to predict the topic flow within discourse. Initially, familiar topics and speakers may be used in quiet conditions with visual cues provided, and then the discourse material can be increased in complexity, unfamiliar and multiple speakers can be introduced, as well as noise and visual distractions. As a result, the patient becomes better able to fill information gaps when individual sound segments or even entire words and phrases are misperceived. With this type of training approach, patients usually receive counseling on how to manipulate context so that they reduce listening difficulty, and how to recover if a predictive or perceptual error results in a communication breakdown.

Auditory training is not routinely used with all adults with hearing loss but tends to be reserved for individuals who have sustained a recent change in auditory status or a substantive increase in auditory demands. For example, adults with sudden deafness, recent cochlear implant recipients, people switching to dramatically different hearing aids with different signal-processing schemes, students entering college, or individuals who are beginning a new job that is auditorily demanding might benefit from auditory training (see cochlear implants; evaluation of cochlear implant candidacy in adults; auditory brainstem implant). Patients who do not make expected improvements in audition and speech after the fitting of a hearing aid or sensory implant also are reasonable candidates for auditory training (see speech perception indices). The fact that most adults do not elect to receive auditory training and typically are not referred for auditory training may be a consequence of the limited data documenting the effectiveness and efficacy of auditory training programs. Only a small number of studies have been published that have assessed auditory training outcomes in adults with hearing loss. Rubenstein and Boothroyd (1987) found only modest benefit with sentence- and syllable-level auditory training with adults who had been successful hearing aid wearers, but they did observe maintenance of gains that were obtained. Walden et al. (1981) found that adults newly fitted with hearing aids benefited significantly from systematic consonant discrimination training, while Kricos and Holmes (1996) found that older adults with previous hearing aid experience did not benefit from consonant and vowel discrimination training but did benefit from active listening training. Auditory training usually focuses on speech and language stimuli, but music perceptual training programs have been developed for cochlear implant recipients and appear to be effective (Gfeller et al., 1999). In addition, auditory training is more strongly advocated for infants and children with hearing loss than for adults, but even fewer interpretable studies have been reported to support its application in children and infants.

Although supporting literature is limited with respect to auditory training of hearing-impaired populations, perceptual training studies with normal-hearing individuals suggest that the impact of auditory training on perception may be underestimated. For example, normal-hearing adults and children have been trained to perceive non-native speech contrasts (Werker and Tees, 1984; Bradlow et al., 1997; Wang et al., 1999). Although not all speech contrasts can be learned equally well, and adults usually fail to reach native speaker performance levels, the effects of training are retained over months and show generalization within and across sound categories (McClaskey, Pisoni, and Carrell, 1983; Lively et al., 1994; Tremblay et al., 1997). Digitally manipulating specific acoustic parameters of speech does not always improve speech perception in expected ways, but shaping speech perception by gradually adjusting more difficult acoustic properties is under investigation in various disordered populations and may prove fruitful in the future for persons with hearing loss (Bradlow et al., 1999; Habib et al., 1999; Merzenich et al., 1996; Thibodeau, Friel-Patti, and Britt, 2001). Furthermore, Kraus and colleagues (Kraus et al., 1995; Tremblay et al., 1997) have argued that auditory training impacts the physiology of the central auditory system and might result in cortical and subcortical reorganization. If neural reorganization occurs after the fitting of a hearing aid or cochlear implant, then it is likely that these types of patients would be sensitive to intensive auditory training during the reorganization period (Kraus, 2001; Purdy, Kelly, and Thorne, 2001).

See also auditory brainstem implant; speech disorders secondary to hearing impairment acquired in adulthood; speech tracking; speechreading training and visual tracking.

—Sheila Pratt

References


Further Readings


Classroom Acoustics

The acoustic environment of a classroom is a critical factor in the psychoeducational and psychosocial development of children with normal hearing and children with hearing impairment. Inappropriate levels of classroom reverberation, noise, or both can deleteriously affect speech perception, reading and spelling ability, classroom behavior, attention, concentration, and academic achievement. Poor classroom acoustics have also been shown to negatively affect teacher performance (Crandell, Smaldino, and Flexer, 1995). This article discusses several acoustic factors that can compromise communication between teacher and child. These acoustic factors include (1) the level of the background noise in the classroom, (2) the relative intensity of the information-carrying components of the speech signal to a non-information-carrying signal or noise (signal-to-noise ratio, S/N), (3) the degree to which the temporal aspects of the information-carrying components of the speech signal are degraded (reverberation), and (4) the distance from the speaker to the listener.

Background Noise

Background noise refers to any auditory disturbance that interferes with what a listener wants or needs to hear. Background noise levels in classrooms are often too high for children to perceive speech accurately. In general, background classroom noise affects a child’s speech recognition by reducing, or masking, the highly redundant acoustic and linguistic cues available in the teacher’s voice. Because the energy of consonant phonemes is considerably less than that of vowel phonemes, background noise in a classroom often masks the consonants more than the vowels. Loss of consonant information has a great effect on speech recognition because the vast majority of the cues important for accurate speech recognition are carried by the consonants.

In most listening environments, the fundamental determinant for speech recognition is not the overall level of the room noise, but rather the relationship between the intensity of the signal and the intensity of the background noise at the listener’s ear. This relationship is referred to as the signal-to-noise ratio (S/N). Speech recognition ability tends to be highest at favorable S/Ns and decreases as the S/N of the listening environment is reduced. Speech recognition in adults with normal hearing is not severely degraded until the S/N approximates 0 dB (speech and noise are at equal intensities). The speech recognition performance of children with sensorineural hearing loss (SNHL), however, is reduced in noise when compared with the performance of children with normal hearing (Finitzo-Hieber and Tillman, 1978). Specifically, children with SNHL require the S/N to be improved by 4–12 dB, and by an additional 3–6 dB in rooms with moderate levels of reverberation, for them to obtain recognition scores equal to those of normal hearers (Crandell and Smaldino, 2000, 2001). While it is recognized that listeners with SNHL experience greater speech recognition deficits in noise than normal hearers, a number of populations of children with “normal hearing” sensitivity also experience significant difficulties recognizing speech in noise. These populations of normal-hearing children include young children (less than 15 years old); those with conductive hearing loss; children with a history of recurrent otitis media; those with a language disorder, articulation disorder, dyslexia, or learning disability; non-native speakers of English, and others with various degrees of hearing impairment or developmental delays. Due to high background noise levels, the range of S/Ns in classrooms has been reported to be between approximately −7 and +5 dB.

Reverberation

Reverberation refers to the prolongation or persistence of sound within an enclosure as sound waves reflect off hard surfaces (bare walls, ceilings, windows, floor) in the room. Operationally, reverberation time (RT) refers to the amount of time it takes for a sound, at a specific frequency, to decay 60 dB (or one-millionth of its original intensity) following termination of the signal. Excessive reverberation degrades speech recognition through the masking of direct and early-reflected energy by reverberant energy. Generally speaking, speech recognition tends to decrease as the RT of the environment increases. Speech recognition in individuals with normal hearing is often not compromised until the RT exceeds approximately 1.0 s (Nabelek and Pickett, 1974a, 1974b). Listeners with SNHL, however, need considerably shorter RTs (0.4–0.5 s) for maximum speech recognition (Crandell, 1991; Crandell, Smaldino, and Flexer, 1995). Studies have also indicated that the populations of “normal-hearing” children discussed previously also have greater speech recognition difficulties in reverberation than do young adults with normal hearing. Unfortunately, RTs for classrooms are often reported to range from 0.4 to 1.2 s.

Effects of Noise and Reverberation on Speech Recognition

In all educational settings, noise and reverberation combine in a synergistic manner to adversely affect speech recognition. That is, noise and reverberation affect speech recognition more than would be expected from an examination of their individual effects on speech perception. It appears that this occurs because reverberation fills in the temporal gaps in the noise, making the noise more steady state in nature and a more effective masker. As with noise and reverberation in isolation, research indicates that listeners with hearing impairment and “normal-hearing” children experience greater speech recognition difficulties in noise and reverberation than adult normal listeners (see Crandell, Smaldino, and Flexer, 1995).
Speaker–Listener Distance

In classrooms, the acoustics of a teacher’s speech signal changes as it travels to the child. The direct sound is that sound which travels from the teacher to a child without striking other surfaces within the classroom. The power of the direct sound decreases with distance because the acoustic energy spreads over a larger area as it travels from the source. Specifically, the direct sound decreases 6 dB in SPL with every doubling of distance from the source sound. This phenomenon, called the inverse square law, occurs because of the geometric divergence of sound from the source. Because the direct sound energy decreases so quickly, only those listeners who are seated close to the speaker will be hearing direct sound energy. At slightly farther distances from the speaker, early sound reflections will reach the listener. Early sound reflections are those sound waves that arrive at a listener within very short time periods (approximately 50 ms) after the arrival of the direct sound. Early sound reflections are often combined with the direct sound and may actually increase the perceived loudness of the sound. This increase in loudness can improve speech recognition in listeners with normal hearing. As a listener moves farther away from a speaker, reverberation dominates the listening environment. As sound waves strike multiple classroom surfaces, they generally decrease in loudness in the increased path length they travel as well as the partial absorption that occurs at each reflection off the classroom surfaces. Some reverberation is necessary to reinforce the direct sound and to enrich the quality of the sound. However, reverberation can lead to acoustic distortions of the speech signal, including temporal smearing and masking of important perceptual cues.

Speech recognition in classrooms depends on the distance of the child from the teacher. If the child is within the critical distance of the classroom (the point at which the intensity of the direct sound and reflected sound are equal in intensity), reflected sound waves have minimal effects on speech recognition. The critical distance in many classrooms is often 5–6 feet from the teacher. Beyond the critical distance, however, the reflections can compromise speech recognition if there is enough of a spectrum or intensity change in the reflected sound to interfere with the recognition of the direct sound. Overall, speech recognition scores tend to decrease until the critical distance of the classroom is reached. Beyond the critical distance, recognition ability tends to remain essentially constant unless the classroom is very large (such as an auditorium). In such environments, speech recognition may continue to decrease as a function of increased distance. These findings suggest that speech recognition ability can be maximized by decreasing the distance between a speaker and listener only within the critical distance of the classroom. Thus, preferential seating, while not a bad idea to improve visual perception, may only minimally improve auditory speech perception.

Acoustic Modifications of the Classroom

The fundamental strategy for improving speech perception within a classroom is acoustic modification of that environment. The most effective procedure for achieving this goal is through appropriate planning with contractors, school officials, architects, acoustical consultants, audiologists, and teachers for the hearing impaired before the design and construction of the building. Acoustic guidelines for hearing-impaired and “normal-hearing” populations indicate the following: (1) S/Ns should exceed +15 dB, (2) unoccupied noise levels should not exceed 30–35 dBA, and (3) RTs should not surpass 0.4–0.6 s (ASHA, 1995). Unfortunately, such guidelines are rarely achieved in most listening environments. Crandell and Smaldino (1995) reported that none of the 32 classrooms in which they measured sound levels met recommended criteria for noise levels. Only 27% of the classroom met criteria for reverberation. A new American National Standards Institute (ANSI S12.60) document entitled “Acoustical Performance Criteria, Design Requirements and Guidelines for Schools” (ANSI, 2002) promises to set a national standard for acoustics in classrooms where learning occurs. If acoustic modifications to learning spaces cannot reduce noise and reverberation to appropriate levels, then hearing assistive technologies, such as frequency modulation (FM) systems, should be implemented for children with “normal hearing” or hearing impairment.

—Carl C. Crandell and Joseph J. Smaldino

References


**Further Readings**


**Clinical Decision Analysis**

Clinical decision analysis (CDA) is a quantitative strategy for making clinical decisions. The techniques of CDA are largely derived from the theory of signal detection, which is concerned with extracting signals from noise. The theory of signal detection has been used to study the detection of auditory signals by the human observer. This article provides a brief overview of CDA. More comprehensive discussions of CDA and its application to audiological tests can be found in Turner, Robinette, and Bauch (1999), Robinette (1994), and Hyde, Davidson, and Alberti (1990).

Assume that we are using a clinical test to distinguish between two conditions, such as disease versus no disease or hearing loss versus normal hearing. Most tests would produce a range of scores for each condition and therefore could be thought to contain some “noise” in their results. More important, there may be an overlap in the scores produced by a test for each condition, creating the potential for error. That is, a particular score could, with finite probability, be produced by either condition. CDA is well-suited to deal with this type of problem.

For the discussions and examples in this article, we will assume that we are trying to identify hearing loss. Of course, CDA can be used with diagnostic tests that are designed to identify a variety of diseases and conditions.

**Decision Matrix**

A patient is given a test, and the outcome of the test is either positive or negative for hearing loss. The test is fallible and makes errors, reflecting the “noise” in the testing process. There are four possible outcomes, determined by the test result and the hearing of the patient. These outcomes are represented by a $2 \times 2$ decision matrix (Fig. 1). If the patient has hearing loss, a positive test is called a hit and a negative result a miss. If the patient has normal hearing, a positive result is a false alarm and a negative result is a correct rejection. Different terminology is sometimes used. A hit is a true positive; a miss is a false negative; a false alarm is a false positive; a correct rejection is a true negative.

The elements of the matrix in Figure 1 represent the number of hits, misses, false alarms, and correct rejections when the test is given to a number of patients. The hits and correct rejections are correct decisions, whereas misses and false alarms are errors. A basic property is that the number of hits plus misses always equals the number of patients with the condition to be identified, e.g., hearing loss. The number of false alarms plus correct rejections always equals the number of patients without the condition, e.g., normal hearing.

**Hit Rate, etc.**

The number of hits, misses, false alarms, and correct rejections can be used to calculate a variety of measures of test performance. The most basic measures are hit rate (HT), miss rate (MS), false alarm rate (FA), and correct rejection rate (CR). Hit rate (also called true positive rate and sensitivity) is the percentage of hearing loss patients correctly identified as positive for hearing loss; miss rate (also called false negative rate) is the percentage of hearing loss patients incorrectly identified as negative. False alarm rate (also called false positive rate) is

<table>
<thead>
<tr>
<th>HEARING STATUS</th>
<th>TEST RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing Loss (pl)</td>
<td>Hit (ht)</td>
</tr>
<tr>
<td>Normal Hearing (pn)</td>
<td>False Alarm (fa)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing Loss</td>
<td>Hit (ht)</td>
<td>Miss (ms)</td>
</tr>
<tr>
<td>Normal Hearing</td>
<td>False Alarm (fa)</td>
<td>Correct Rejection (cr)</td>
</tr>
</tbody>
</table>
the percentage of normal-hearing patients incorrectly called positive; correct rejection rate (also called true negative rate and specificity) is the percentage of normal-hearing patients correctly identified as negative. These measures are calculated by the following equations:

\[
HT = \frac{ht}{pl} \\
MS = \frac{ms}{pl} \\
FA = \frac{fa}{pn} \\
CR = \frac{cr}{pn}
\]

where \(pl\) = number of patients with hearing loss, \(pn\) = number of patients with normal hearing, \(ht\) = number of hits, \(ms\) = number of misses, \(fa\) = number of false alarms, and \(cr\) = number of correct rejections. The above equations yield a decimal value \(\leq 1.0\), which can be converted to a percent. In all calculations, the decimal form is used. While all four measures can be calculated, only two, usually \(HT\) and \(FA\), need be considered. This is because \(HT + MS = 100\%\) and \(FA + CR = 100\%;\) thus, \(MS\) and \(CR\) can always be determined from \(HT\) and \(FA\).

The measures of test performance, \(HT\), \(FA\), \(MS\), and \(CR\), can also be expressed as probabilities. \(HT\) is the probability of a positive test result if the patient has hearing loss (Pr[\(+\)/L]); \(FA\) is the probability of a positive result if the patient has normal hearing (Pr[\(+\)/N]); \(MS\) is the probability of a negative result given hearing loss (Pr[\(-\)/L]), and \(CR\) is the probability of a negative result given normal hearing (Pr[\(-\)/N]).

**Probability Distribution Curve**

Consider a theoretical test that produces a score from 0 to 100. The test is administered to two groups of patients, one with hearing loss and one with normal hearing. Each group will produce a range of scores on the test. We could plot a histogram of scores for each population, i.e., the number of patients in a group with a score between 0 and 10, 10 and 20, etc. Next, we divide the number of patients in each score range by the total number of patients in the group to obtain the probability distribution curve (PDC). Essentially, the PDC gives the probability of obtaining a particular score, or range of scores, for each group of patients. PDCs for a theoretical test are shown in Figure 2. Note that there are two PDCs, one for hearing loss and one for normal hearing, and that the two distributions are different.

Because the two PDCs in Figure 2 do not completely overlap, we may use this test to identify hearing loss.
First, however, we must establish a criterion to determine if a test score is positive or negative for hearing loss. In Figure 2, we set the criterion at 50. Normal-hearing patients have, on average, higher scores than hearing loss patients; therefore, a score greater than 50 is negative and a score less than or equal to 50 is positive for hearing loss. Because there is some overlap in the two PDCs, the criterion divides the two PDCs into four regions corresponding to hits, misses, false alarms, and correct rejections. Below the criterion (≤50), hearing loss patients constitute the hits and the normal-hearing patients constitute the false alarms. Above the criterion, the hearing loss patients are the misses and the normal-hearing patients are the correct rejections. Hit rate is the total probability that hearing loss patients will have a test score ≤ 50. This equals the sum of the probabilities for all of the bars below the criterion that correspond to hearing loss patients.

We can select any criterion for a test, but different criteria will produce different test performance. If the criterion is increased (e.g., from 50 to 60), there is an increase in HT, which is good, but there is also an increase in FA, which is bad. If the criterion is reduced (e.g., from 50 to 40), the FA will be reduced, which is good, but so will the HT, which is bad. Thus, we see that there is usually a trade-off between HT and FA when we adjust the criterion.

Another interesting result occurs with extreme criteria. We could set the criterion at 100 and call all results positive for hearing loss. This would produce an HT of 100%; however, FA would also equal 100%. Likewise, with a criterion of 0, FA = 0% but also HT = 0%. Thus, HT and FA can be manipulated by changing the criterion, but the trade-off between HT and FA limits the value of this strategy. Because any HT or FA can be obtained by adjusting the criterion, both HT and FA are needed to evaluate the performance of a test.

The ability of a test to distinguish patients is related to the amount of overlap of the PDCs. If in Figure 2 the two PDCs completely overlapped, then the test would be useless. For any criterion, we would have HT = FA and MS = CR. If there was absolutely no overlap in the PDCs, then the test would be perfect. It would be possible to set the criterion such that HT = 100% and FA = 0%.

Because HT and FA vary significantly with the criterion, we can visualize this relationship using a receiver operating characteristic (ROC) curve, which is a plot of HT versus FA for different criteria. The HT/FA data from Figure 2 are plotted in Figure 3. The shape of the ROC curve is determined by the PDCs.

**Posterior Probabilities**

Consider this situation. We are testing a patient in the clinic, and the test result is positive. We know the hit rate of the test, but hit rate is the probability of a positive result given hearing loss. We do not know if the patient has hearing loss, but we do know the test result. Hit rate tells us little about the accuracy of the test result. What we want is not hit rate, the probability of a positive result given hearing loss (Pr[+L]), but the probability that the patient has hearing loss given a positive test result (Pr[L/+]). This probability is called a posterior probability. Another posterior probability is Pr[N/−]; this is the probability of normal hearing given a negative test result. These two posterior probabilities are important because they are the probability of being correct given a particular test result. There are two other posterior probabilities, Pr[L/−] and Pr[N/+], which are the probability of being incorrect given a test result. The posterior probabilities have been given other names: predictive value and information content. These measures are identical to the posterior probabilities and thus provide the same information.

Because Pr[L/+] + Pr[N/+] = 100% and Pr[L/−] + Pr[N/−] = 100%, we need calculate only two of the four posterior probabilities. To calculate the posterior probabilities we need HT and FA for the test, plus the prevalence (PD) of the disease or condition in the test population. Prevalence is the percentage of the test population that has the disease or condition (e.g., hearing loss) at the time of testing. The posterior probabilities are calculated by the following equations:

\[
Pr[L/+] = \frac{1}{1 + \frac{(FA)(1-PD)}{(HT)(PD)}}
\]

\[
Pr[N/-] = \frac{1}{1 + \frac{(1-HT)(PD)}{(1-FA)(1-PD)}}
\]
The posterior probabilities can also be calculated from the number of hits, misses, false alarms, and correct rejections. Sometimes this is an easier strategy than using the equations above:

\[
\Pr[L/+] = \frac{ht}{ht + fa} \\
\Pr[N/-] = \frac{cr}{cr + ms}
\]

The probability of being correct with a positive test result is the number of hits divided by the total number of positive test results (hits plus false alarms). Likewise, the probability of being correct with a negative test result is the number of correct rejections divided by the total number of negative results (correct rejections plus misses).

When the prevalence of a condition (e.g., hearing loss) is small, the probability of a positive result being correct (Pr(L/+)) is small, even for a test with high HT and small FA. For example, consider a test with HT/FA = 99/5%. Even with this test, which is better than any audiological test, the probability of a positive result being correct is only 29% for PD = 2%. There would be 2.5 false alarms for each hit. When prevalence is small, we should expect more false alarms than hits.

Now consider Pr[N/-], the probability of being correct with a negative test result. When prevalence is low, Pr[N/-] is very large, 99+% for the example above. Only when prevalence is high is there a significant variation in Pr[N/-] with test performance.

**Efficiency**

HT, FA, and PD can also be used to calculate efficiency (EF). EF is the percentage of total test results that are correct and is calculated by

\[
EF = HT \times PD + (1 - FA) \times (1 - PD)
\]

Like the posterior probabilities, efficiency is a function of disease prevalence. When prevalence is small, the false alarm rate drives efficiency more than the hit rate. Thus, a test with a poor HT and a small FA could have a higher EF than a test with a high HT and a modest FA. Because of this, EF is not always a useful measure.

**Conclusion**

Clinical decision analysis provides us with a variety of measures of test performance. These measures are useful for evaluating different tests and for understanding the limitations of a particular test. In general, these measures of performance are not sufficient to identify the “best” test. To determine the best test for a particular application, we may need to also consider other issues, such as the cost or morbidity of the test. The decision as to the best test is based on a cost-benefit analysis, not simply measures of test performance. Nevertheless, these measures of performance are essential to execute a cost-benefit analysis.

—Robert G. Turner

### References


### Cochlear Implants

A cochlear implant is a surgically implantable device that provides hearing sensation to individuals with severe to profound hearing loss who cannot benefit from conventional hearing aids. By electrically stimulating the auditory nerve directly, a cochlear implant bypasses damaged or undeveloped sensory structures in the cochlea, thereby providing usable information about sound to the central auditory nervous system.

Although it has been known since the late 1700s that electrical stimulation can produce hearing sensations (see Simmons, 1966), it was not until the 1950s that the potential for true speech understanding was demonstrated. Clinical applications of cochlear implants were pioneered by research centers in the United States, Europe, and Australia. By the 1980s, cochlear implants had become a clinical reality, providing safe and effective speech perception benefit to adults with profound hearing impairment. Since that time, the devices have become more sophisticated, and the population that can benefit from implants has expanded to include children as well as adults with some residual hearing sensitivity (Wilson, 1993; Shannon, 1996; Loizou, 1998; Osberger and Koch, 2000).

The function of a cochlear implant is to provide hearing sensation to individuals with severe to profound hearing impairment. Typically, people with that level of impairment have absent or malfunctioning sensory cells in the cochlea. In a normal ear, sound energy is converted to mechanical energy by the middle ear, and the mechanical energy is then converted to mechanical fluid motion in the cochlea. Within the cochlea, the sensory cells—the inner and outer hair cells—are sensitive transducers that convert that mechanical fluid motion into electrical impulses in the auditory nerve. Cochlear implants are designed to substitute for the function of the middle ear, cochlear mechanical motion, and sensory cells, transforming sound energy into the electrical energy that will initiate impulses in the auditory nerve.

All cochlear implant systems comprise both internal and external components (Fig. 1). Sound enters the microphone, and the signal is then sent to the speech processor, which manipulates and converts the acoustic signal into a special code (i.e., speech-processing...
strategy). The transmitter, which is located inside the headpiece, sends the coded electrical signal to the internal components. The internal device contains the receiver, which decodes the signal from the speech processor, and an electrode array, which stimulates the cochlea with electrical current. The implanted electronics are encased in one of two biocompatible materials, titanium or ceramic. The entire system is powered by batteries located in the speech processor, which is worn on the body or behind the ear.

The transmission link enables information to be sent from the external parts of the implant system to the implanted components. For all current systems, the connection is made through transcutaneous inductive coupling of radio-frequency (RF) signals. In this scheme, an RF carrier signal—in which the important code is embedded—is sent across the skin to the receiver. The receiver extracts the embedded code and determines the stimulation pattern for the electrodes. Most cochlear implant systems also employ back telemetry, which allows the internal components to send information back to the external speech processor to assess the function of the implanted electronics and electrode array.

The first cochlear implants consisted of a single electrode, but since the mid-1980s, nearly all devices have multiple electrodes contained in an array. Typically, cochlear implant electrodes are inserted longitudinally into the scala tympani of the cochlea to take potential advantage of the place-to-frequency coding mechanism used by the normal cochlea. Information about low-frequency sound is sent to electrodes at the apical end of the array, whereas information about high-frequency sounds is sent to electrodes nearer the base of the cochlea. The ability to take advantage of the place-frequency code is limited by the number and pattern of surviving auditory neurons in an impaired ear. Unfortunately, attempts to quantify neuronal survival with electrophysiologic or radiographic procedures before implantation have been unsuccessful (Abbas, 1993).

The first multi-electrode arrays were straight and thin (22 electrode bands on the 25-mm-long array) to minimize the occupied space within the scala tympani (Clark et al., 1983). Advances in electrode technology have led to the development of precurved, spiral-shaped arrays to follow the shape of the scala tympani, allowing the contacts to sit close to the target neurons (Fayad, Luxford, and Linthicum, 2000; Tykocinski et al., 2002). The advantages of the precurved array are an increase in spatial selectivity, a reduction in channel interaction, and a reduction in the current required to reach threshold and comfortable listening levels. In addition, the electrode contacts are oriented toward the spiral ganglion cells, and a “positioner,” a thin piece of Silastic, can be inserted behind the array to achieve even greater spatial selectivity and improve speech recognition performance (Zwolan et al., 2001).

For all systems, electrical current is passed between an active electrode and an indifferent electrode. If the active and indifferent electrodes are remote, the stimulation is termed monopolar. When the active and indifferent electrodes are close to each other, the stimulation is referred to as bipolar. Bipolar stimulation focuses the current within a restricted area and presumably stimulates a small localized population of auditory nerve fibers (Merzenich and White, 1977; van den Honert and Stypulkowski, 1987). Monopolar stimulation, on the other hand, spreads current over a wider area and a larger population of neurons. Less current is required to achieve adequate loudness levels with monopolar stimulation; more current is required for bipolar stimulation. The use of monopolar or bipolar stimulation is determined by the speech-processing strategy and each individual’s response to electrical stimulation.

Two types of stimulation are currently used in cochlear implants, analog and pulsatile. Analog stimulation consists of electrical current that varies continuously in time. Pulsatile stimulation consists of trains of square-wave biphasic pulses. The pattern of stimulation can be either simultaneous or nonsimultaneous (sequential). With simultaneous stimulation, more than one electrode is stimulated at the same time. With nonsimultaneous stimulation, electrodes are stimulated in a specified sequence, one at a time. Typically, analog stimulation is simultaneous and pulsatile stimulation is sequential.

To represent speech faithfully, the coding strategy must reflect three parameters in its electrical stimulation code: frequency, amplitude, and time. Frequency (pitch) information is conveyed by the site of stimulation, amplitude (loudness) is encoded by the amplitude of the
stimulus current, and temporal cues are conveyed by the rate and pattern of stimulation. The first multichannel devices extracted limited information from the acoustic input signal (Millar, Tong, and Clark, 1984). Advances in signal-processing technology have led to the development of more sophisticated processing schemes. One type of strategy is referred to as “n of m,” in which a specified number of electrodes out of a maximum number available are stimulated (Seligman and McDermott, 1995). With this type of processing, the incoming sound is analyzed to identify the filters (frequency regions) with the greatest amount of energy, and then a subset of filters is selected and the corresponding electrodes are stimulated.

In another approach, referred to as continuous interleaved sampling (CIS), trains of biphasic pulses are delivered to the electrodes in an interleaved or non-overlapping fashion to minimize electrical field interactions between stimulated electrodes (Wilson et al., 1991). The amplitudes of the pulses delivered to each electrode are derived by modulating them with the envelopes of the corresponding bandpassed waveforms. With analog stimulation, the incoming sound is separated into different frequency bands, compressed, and delivered to all electrodes simultaneously (Eddington, 1980). In the most recent implementation of this type of processing, the digitized signal is transmitted to the receiver; then, following digital-to-analog conversion, the analog waveforms are sent simultaneously to all electrodes (Kessler, 1999). Bipolar electrode coupling is typically used to limit the area over which electrical current spreads to reduce channel interaction, which is further reduced with the use of spiral-shaped electrodes.

Cochlear implant candidacy is determined only after comprehensive evaluations by a team of highly skilled professionals (see COCHLEAR IMPLANTS IN ADULTS: CANDIDACY). The surgery is performed under general anesthesia and requires about 1–2 hours, either as an inpatient or outpatient procedure. Approximately 4 weeks following surgery, the individual returns to the clinic to be fitted with the external components of the system. Electrical threshold and most comfortable listening levels are determined for each electrode, and other psychophysical parameters of the speech-processing scheme are programmed into the speech processor. Multiple visits to the implant center are necessary during the first months of implant use as the individual grows accustomed to sound and as tolerance for loudness increases.

Most adults who acquire a severe to profound hearing loss after language is acquired (postlingual hearing loss) demonstrate dramatic improvements in speech understanding after relatively limited implant experience (Fig. 2). Improvements in technology have led to incremental improvements in benefit, which in turn have led to expanded inclusion criteria (Skinner et al., 1994; Osberger and Fisher, 1999). There are large individual differences in outcome, and although there is no reliable method to predict postimplant performance, age at onset of significant hearing loss, duration of the loss, and degree of preoperative residual hearing significantly affect speech recognition abilities (Tyler and Summerfield, 1996; Rubinstein et al., 1999). Many adults are able to converse on the telephone, and cochlear implants can improve the quality of life (Knutson et al., 1998). Adults with congenital or early-acquired deafness and children (see COCHLEAR IMPLANTS IN CHILDREN) also derive substantial benefit from cochlear implants.

Technological advances will continue, with higher processing speeds offering the potential to stimulate the auditory nerve fibers in a manner that more closely approximates that of normal hearing. Studies are under way to evaluate the benefit of bilateral implants (Gantz et al., 2002). New miniaturization processes will result in smaller behind-the-ear processors and, eventually, a fully implantable system with rechargeable battery technology. In the early days of cochlear implants, few people realized that this technology would become the most successful of all prostheses of the central nervous system.

—Mary Joe Osberger
References


Further Readings


Cochlear Implants in Adults: Candidacy

There are many potential advantages to measuring the benefit obtained from cochlear implants. These include:

- Determining selection criteria
- Setting the cochlear implant: selecting and modifying programming options
- Monitoring performance

The most important reason for the evaluation of cochlear implants is in the selection criteria process. Specifically, speech perception tests are critical to determine whether a hearing aid user might do better with a cochlear implant. Here we focus on this subject and review some of the tests used in evaluation.

Selection Criteria

Guidelines traditionally have depended on how much benefit is obtained from hearing aids, and how much
Table 1. Principles Involved When Considering Monaural Implantation

<table>
<thead>
<tr>
<th>Binaural Test Results</th>
<th>Monaural Implant Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>If binaural scores are high, and both ears are contributing equally</td>
<td>Do not implant.</td>
</tr>
<tr>
<td>If binaural scores are high, and one ear is not contributing</td>
<td>Consider implanting poorer ear to improve spatial hearing.</td>
</tr>
<tr>
<td>If binaural scores are medium, and both ears are contributing equally</td>
<td>Implant ear with shorter duration and/or better thresholds.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is contributing less than the other, and best monaural = binaural (no binaural benefit)</td>
<td>Implant poorer ear, to improve spatial hearing.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is contributing less than the other, and best monaural &lt; binaural (is binaural benefit)</td>
<td>Implant poorer ear, to improve binaural benefit and preserve better ear.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is not contributing, and best monaural = binaural (no binaural benefit)</td>
<td>Implant better ear, to improve speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are poor, with one ear contributing more than the other</td>
<td>Implant poorer ear, to improve spatial hearing and speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are poor, with ears contributing equally</td>
<td>Implant better ear, to improve speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are high, and both ears are contributing equal</td>
<td>Implant ear with shorter duration and/or better thresholds.</td>
</tr>
</tbody>
</table>

benefit might be expected from a cochlear implant. This is difficult, because there are limited databases with such information. Specific guidelines for selection criteria change regularly, are influenced by company and clinic protocols and depend on whether the device is investigational. Generally, the “best aided” performance (with appropriately fit binaural hearing aids) is used to determine if an implant is desirable. When considering a monaural implant, some centers implant the poorer ear to “save” the better ear or to allow the patient to use a cochlear implant in one ear and a hearing aid (better ear) in the other. At other centers, however, the implant is placed in the better ear to maximize implant performance in an ear with more nerve fibers.

First, we shall discuss general guidelines. We refer to poor, medium, and high speech perception scores, realizing that this is arbitrary. The actual values will depend on the test, and will change as overall implant performance improves. The principles we are promoting involve binaural testing and determining the contribution from each ear individually.

Table 1 lists the principles involved when considering monaural implantation. Speech perception tests are conducted typically with the patient using hearing aids testing the right, left, and binaural conditions. The results from each individual ear can then be compared and the best monaural condition can be determined. In addition, the best monaural condition can be compared to the binaural condition to determine if there is a binaural advantage. Table 2 lists the principles involved in considering binaural implantation.

Protocols for Evaluation

This section reviews some of the tests used in the evaluation process.

Isolated Words. The presentation of isolated words (Peterson and Lehiste, 1962; Tillman and Carhart, 1966) has the advantage that linguistic and cognitive factors are minimized, and clinicians are familiar with the tests. When isolated word lists are used, the vocabulary should be common words.

Ongoing Speech. Sentence perception includes the processing of information at a more rapid, natural rate, compared to the presentation of isolated words (Silverman and Hirsh, 1955; Boothroyd, Hanin, and Hnath, 1985; Nilsson, Soli, and Sullivan, 1994). Performance can be affected by memory, learning, and familiarity with items as a result of repeated use of the test lists. In addition, a patient who is more willing to guess or who is better at using contextual clues may score higher than someone who has similar speech perception abilities but is less willing to guess. Therefore, it is important that sentence length be short, lots of sentences be available, and the test-retest reliability of the materials be known.

Speech Perception Testing in Noise. Many realistic listening situations involve background noise, resulting in differences in speech perception that are not apparent when testing is performed in quiet. Background noise can result in a “floor effect” (near 0% correct). Therefore, in some situations a favorable signal-to-noise ratio (S/N) is selected individually, or testing that adaptively varies the S/N is used (Plomp and Mimpfen, 1979; Levitt, 1992).

Subjective Ratings. Quality ratings measure a more global attribute of speech. For example, ratings might include the ease of listening or clarity of sound. We have developed a quality rating test that includes realistic listening situations. Patients are asked to listen to each sound and to rate it on a scale from 0 (unclear) to 100 (clear). Figure 1 shows the results from an adult cochlear implant patient comparing a high-rate, roving, n-of-m strategy (Wilson, 2000) to a SPEAK strategy (Skinner
Table 2. Principles Involved When Considering Binaural Implantation

<table>
<thead>
<tr>
<th>Binaural Test Results</th>
<th>Binaural Implant Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>If binaural scores are high, and both ears are contributing</td>
<td>• Do not implant.</td>
</tr>
<tr>
<td>If binaural scores are high, and one ear is not contributing</td>
<td>• Implant poorer ear to improve spatial hearing.</td>
</tr>
<tr>
<td>If binaural scores are medium, and both ears are contributing equally</td>
<td>• Implant binaurally.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is contributing more than the other, and best monaural = binaural (no binaural benefit)</td>
<td>• Do not implant (conservative approach).</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is contributing more than the other, and best monaural &lt; binaural (is binaural benefit)</td>
<td>• Implant poorer ear, to improve spatial hearing.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is not contributing, and best monaural = binaural (no binaural benefit)</td>
<td>• Implant binaurally, to improve spatial hearing and speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is contributing more than the other, and best monaural &lt; binaural (is binaural benefit)</td>
<td>• Do not implant, to preserve current performance levels.</td>
</tr>
<tr>
<td>If binaural scores are medium, one ear is not contributing, and best monaural = binaural (no binaural benefit)</td>
<td>• Implant poorer ear, to improve binaural benefit and to preserve better ear.</td>
</tr>
<tr>
<td>If binaural scores are poor, with one ear contributing more than the other</td>
<td>• Implant poorer ear, to improve speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are poor, with ears contributing equally</td>
<td>• Implant poorer ear, to improve speech in quiet.</td>
</tr>
<tr>
<td>If binaural scores are poor, with one ear contributing more than the other</td>
<td>• Implant binaurally, to improve speech in quiet and spatial hearing.</td>
</tr>
<tr>
<td>If binaural scores are poor, with ears contributing equally</td>
<td>• Implant binaurally.</td>
</tr>
</tbody>
</table>

Figure 1. Average subjective ratings for a high rate, roving, n-of-m strategy, worn bilaterally, compared to bilateral SPEAK.

Figure 2. Response form from the Audiovisual Feature Test. The patient hears one of the ten items and must point to the item that was presented.
et al., 1997). The test was sufficiently sensitive to show a strategy preference for all sounds. Information such as this can be extremely helpful when selecting and modifying programming options.

**Speech Features.** It is advantageous to determine what kinds of speech sounds are perceived to understand which features are being transmitted by the cochlear implant, and how to focus rehabilitation. We use a variety of consonant and vowel tests. The Iowa Medial Consonant Test (Tyler, Preece, and Lowder, 1983) presents items in an “ee/C/ee” (13-choice version) or “aa/C/aa” (16- and 24-choice versions) context, where /C/ represents a variety of consonants (e.g., “aa/P/aa,” “aa/M/aa,” “aa/D/aa,” etc.).

For adult patients who are poor performers and for children, the Audiovisual Speech Feature Test is easier and can be used (Tyler, Fryauf-Bertschy, and Kelsay, 1991). Figure 2 shows the items included in the test.

**Spatial Hearing.** Another attempt to make testing more realistic includes measurement of spatially separate speech and noise and the localization of sounds (Shaw, 1974; Middlebrooks and Green, 1991; Wightman and Kistler, 1997). More individuals are being fitted with either binaural implants or have a hearing aid and a cochlear implant. We typically measure speech from the front and noise from the right or left. This allows for the measurement of the “head shadow” effect and a “binaural squelch effect.”

Localization is based on interaural time, amplitude, and spectral differences between ears (Mills, 1972; Wightman and Kistler, 1992). For binaural cochlear implant recipients the fine details of this information may not be available. To test localization, everyday sounds are presented through 8 loudspeakers. The loudspeakers are arranged in an arc, the patient is asked to indicate which speaker the sound came from (Fig. 3). Figure 4a shows results from one patient who was tested wearing only the right cochlear implant, Figure 4b only the left cochlear implant, and Figure 4c with both implants at the same time.
Conclusions

We have provided a brief overview of some of the more common issues involved in evaluating adult cochlear implant users. The most important task is to determine candidacy. Speech perception tests measuring binaural hearing aid benefit are needed to determine either monaural or binaural cochlear implant candidacy. Several different tests measure a wide range of potential hearing abilities.

—Richard S. Tyler and Shelley Witt

References


Further Readings


Cochlear Implants in Children

The substantial benefit derived from cochlear implants by adults (see COCHLEAR IMPLANTS AND COCHLEAR IMPLANTS IN ADULTS: CANDIDACY) led to the application of these devices in children. Unlike adults, however, most pediatric candidates acquire their deafness before speech and language are learned (prelingually deafened). Thus, children must depend on an auditory prosthesis to learn the auditory code underlying spoken language—a formidable task, given the exquisite temporal and frequency-resolving powers of the normal ear. On the other hand, young children may be the most successful users of implantable auditory prostheses because of the plasticity of the central nervous system.

The challenges faced in determining candidacy in children require balancing the risks of surgery versus the potential benefits of early implantation for the acquisition of spoken language. Initially, the use of cochlear implants in children was highly controversial. Thus, candidacy requirements were stringent, and the first children to receive cochlear implants were older (school-age or adolescents) and demonstrated no benefit from conventional hearing aids—not even sound awareness—even after many years of use and rehabilitation. These children were considered “ideal” cochlear implant candidates because their hearing could be reliably evaluated and it was obvious that no improvement in their auditory skills would occur with conventional hearing aids.

The first devices used with children were single-channel implants (see COCHLEAR IMPLANTS). Even though performance was limited, the children who received these devices derived more benefit from their implants.
than from conventional hearing aids (Thielemeir et al., 1985; Robbins, Renshaw, and Berry, 1991). A small percentage of these children achieved remarkable levels of word recognition through listening alone, although many of them had early acquired deafness with some normal auditory experience prior to the onset of their hearing loss (Berliner et al., 1989). Benefits also were documented in speech production and language acquisition (Osberger, Robbins, Berry, et al., 1991). Clearly, the early pioneering work with single-channel devices demonstrated the safety and effectiveness of implantable auditory prostheses in children and paved the way for the acceptance of cochlear implants as a medical treatment for profound deafness.

Eventually children received multichannel cochlear implants, especially as results indicated superior outcomes with these devices compared with single-channel implants (Osberger, Robbins, Miyamoto, et al., 1991). Since that time, numerous research studies have documented the substantial benefits that children with profound hearing loss obtain from multichannel cochlear implants (see Kirk, 2000; Waltzman, 2000). Numerous speech perception tests have been developed to assess implant candidacy and benefit, even in very young children (Kirk, 2000; Zimmerman-Phillips, Robbins, and Osberger, 2000). A finding common to all studies is the long time course over which children acquire auditory, speech, and language skills, even with multichannel devices (Tyler et al., 2000) (Fig. 1). This is not unexpected, given the number of years required for similar skill acquisition by hard-of-hearing children who use conventional hearing aids.

With continued clinical experience, improvements in technology, and documented benefits, cochlear implants gained greater acceptance, and candidacy criteria were expanded. Children received implants at increasingly younger ages, and it is now common practice to place implants in children as young as 2 years, with a growing trend for children as young as 12 months of age to receive cochlear implants (Waltzman and Cohen, 1998). Identification of hearing loss at an early age has also contributed to implantation in children at increasingly younger ages. Evidence suggests that children receiving implants at a younger age achieve higher levels of performance with their devices than children receiving implants at an older age (Fryauf-Bertschy et al., 1997; Waltzman and Cohen, 1998). Significant differences in postimplant outcome have been documented in children who receive implants before age 3 years. Children who received cochlear implants between ages 12 and 23 months demonstrated better auditory skills after implantation than children who received implants between the ages of 24 and 36 months (Osberger et al., 2002). Thus, a difference of as little as 1 year in age at the time of implantation had a significant impact on the rate of auditory skill development in these young children.

Even though the current trend is to provide implants to children at younger ages, older children continue to receive cochlear implants (Osberger et al., 1998). Some of these children have residual hearing and demonstrate benefit from conventional hearing aids. Implantation is often delayed because it takes longer to determine whether a plateau in auditory development has been reached. In addition, the audiological candidacy criteria were more stringent when these children were younger, and thus they were not considered appropriate candidates because they had too much hearing. Over time, however, audiological criteria in children have been expanded for implants (Zwolan et al., 1997). Following implantation, children with preoperative speech perception abilities demonstrate remarkable auditory recognition skills and achieve higher levels of performance with their implants than they did with hearing aids (Fig. 3).

Other factors besides age influence cochlear implant benefit in children. Communication method also impacts the postimplant performance in children. Most studies have found that children who use oral communication...
Audition, speaking, lipreading achieve higher levels of performance with their implants than do children who are educated using total communication (English-based sign language with audition, speaking, lipreading) (Meyer et al., 1998). The trend for better implant performance in children who use oral communication has been shown in older children (Fig. 4) as well as in very young children (Osberger et al., 2002). This finding indicates that oral education programs more effectively emphasize the use of auditory information provided by an implant than do total communication programs. In fact, since multichannel cochlear implants became available, there has been a dramatic increase in the number of educational programs that employ oral communication, because a greater number of children have the potential to acquire spoken language through audition.

In addition to auditory perceptual benefits, children with cochlear implants show significant improvement in their receptive and expressive language development (see Robbins, 2000). Improvements in the use of communication strategies and conversational skills have also been reported (Tait, 1993; Nicholas, 1994), and more children with implants demonstrate higher levels of reading achievement than reported for their peers with hearing aids (Spencer, Tomblin, and Gantz, 1999). Nonetheless, even with marked improvements in performance, children with cochlear implants remain delayed in linguistic development compared to children of the same chronological age with normal hearing. However, children with cochlear implants do not continue to fall farther behind in their language performance, as has been reported for their profoundly hearing-impaired peers with hearing aids. As deaf children receive implants at younger ages, the gap between their skills and the skills of their age-matched peers with normal hearing will lessen.

Speech production skills also improve after implantation. Studies have shown improved production of segmental and suprasegmental features of speech and overall speech intelligibility (Tobey, Geers, and Brenner, 1994; Robbins et al., 1995). Dramatic improvements in speech production are often apparent after only several months of implant use, even in very young children who had little or no auditory experience prior to implantation. In very young children, improvements in vocalizations are usually the most noticeable changes following implantation (Zimmerman-Phillips, Robbins, and Osberger, 2000).

Cochlear implants are now accepted as an effective treatment for profound deafness. Many profoundly deaf children gain access to the auditory and linguistic code of spoken language with these devices, an accomplishment realized by only a limited number of deaf children with hearing aids. Profoundly deaf children with cochlear implants often function as well as children with less severe hearing impairments with hearing aids (Boothroyd and Eran, 1994; Meyer et al., 1998). Consequently, deaf children with implants acquire spoken language vicariously through incidental learning, requiring fewer special support services in school. Evidence suggests that more deaf children who use implants are being mainstreamed in regular classrooms than their peers with hearing aids (Francis et al., 1999). Thus, the long-term educational costs for children with cochlear implants will be less than for deaf children with hearing aids, resulting in a net savings to society. In addition, cochlear implants have a positive effect on the quality of life of deaf children, and have also been found to be a cost-effective treatment for deafness (Cheng et al., 2000). Whereas the impact of cochlear implants on educational and vocational achievement will take many years to establish, it is clear that these devices have changed the lives of many deaf children and their families.

—Mary Joe Osberger

References


**Further Readings**


Dichotic Listening

Dichotic listening refers to listening to two different signals presented simultaneously through earphones, one signal to the left ear (LE) and a different signal to the right ear (RE). Although the results have been expressed in different ways, the most common approach is to calculate \%RE, \%LE, and the difference score (%RE – %LE). The difference score describes the percentage ear advantage and may be a right ear advantage (REA), left ear advantage (LEA), or no ear advantage (NoEA).

Dichotic listening is a psychophysical process, the testing of which is used to assess certain aspects of central auditory function. The outcomes of experiments have led to the development of ear–brain hypotheses; an REA is accepted as evidence of left hemispheric dominance for processing and an LEA as evidence of right hemispheric dominance. A NoEA is sometimes interpreted to mean that brain dominance has not been well established (Gerber and Goldman, 1971).

When the signals are speech (usually consonant-vowel [CV] nonsense syllables), the commonly reported outcome is an REA, and interpretation of the REA in relation to left hemispheric dominance has been based on four assumptions: (1) ipsilateral auditory pathways are suppressed during dichotic stimulation (Milner, Taylor, and Sperry, 1968); (2) information from each ear arrives at the contralateral hemisphere in equivalent states (Studdert-Kennedy and Shankweiler, 1970); (3) the left hemisphere, which is language dominant for at least 95% of the right-handed population and about 70% of the left-handed population (Penfield and Roberts, 1959; Annett, 1975; Rasmussen and Milner, 1977), is principally responsible for extracting phonetic information from the different signals presented to the RE and LE (Studdert-Kennedy and Shankweiler, 1970; Studdert-Kennedy, Shankweiler, and Pisoni, 1972); and (4) the lower LE score implies “loss of information” during interhemispheric transmission from the right hemisphere to the left hemisphere via the corpus callosum (Studdert-Kennedy and Shankweiler, 1970; Studdert-Kennedy, Shankweiler, and Pisoni, 1972; Berlin et al., 1973; Brady-Wood and Shankweiler, 1973; Cullen et al., 1974; Repp, 1976). Each assumption is critical to the validity of an ear–brain hypothesis.

Distinction Between a True Ear Advantage and an Observed Ear Advantage

A true ear advantage is thought to reflect differences in the transmission capacities of the auditory channels from the RE and LE to a common, centrally located processing area that, for speech, is located in the left hemisphere. An observed ear advantage may arise from at least four sources: a true ear advantage, decision variables, stimulus variables, and measurement error (Speaks, Niccum, and Carney, 1982). Proper counterbalancing of experimental conditions can control stimulus variables, measurement error can be minimized by presenting a sufficient number of listening trials (Repp, 1977; Speaks, Niccum, and Carney, 1982), and the role of decision variables will be addressed subsequently.

Optional Psychophysical Procedures

Kimura (1961) used a recall task. Listeners received three pairs of spoken digits and were asked to recall as many of the six digits as possible. A second procedure—the most commonly used—involves a two-ear recognition task. One pair of signals is presented, the listener is to attend “equally” to both the RE and LE, and the listener provides two responses from a closed message set (Studdert-Kennedy and Shankweiler, 1970; Berlin et al., 1972; Speaks, Niccum, and Carney, 1982). In a variation of the two-ear recognition task, the listener attends to both ears but provides only one response (Repp, 1977). A third procedure employs an ear-monitoring task in which two signals are presented but the listener is asked to attend selectively to one ear and supply only one response (Hayden, Kirsten, and Singh, 1979). Although those three tasks can be used with speech signals, their use for nonspeech signals is more problematic because the listener is required to both recall and name the signals heard.

A fourth technique applies the theory of signal detection (TSD) and involves a yes/no target-monitoring task (Katsuki et al., 1984). The TSD approach can be used with either speech or nonspeech signals, but the details will be described for speech signals. The message set consists of six CV nonsense syllables. On a given test block of 40 dichotic trials, one syllable is designated the target. Contralateral interference is provided by the other five syllables, each appearing equally often. The target is present on only half (20) of the trials within a listening block; hence the a priori probability of target-present trials is 0.50. The listener attends to a designated ear and responds “yes” to vote that the target was present or “no” to vote that the target was not present. Over the course of six listening blocks, each of the six syllables serves as the target, and contralateral interference is provided by the other five syllables. The scores for each ear are expressed by P(C)max, a \textit{d’}-based statistic, the ear advantage is given by P(C)max_{RE} – P(C)max_{LE}, and listener criterion is expressed by \textit{b}. The effects of decision variables on the outcome of the yes/no target-monitoring task are minimized because both hit and
false alarm responses are incorporated in the calculation of \( P(C)_{\text{max}} \).

**Choice of Metric**

When the task is yes/no target monitoring, the ear advantage should be expressed by \( P(C)_{\text{max}} \) for the right ear and \( P(C)_{\text{maxL}} \). For the commonly used two-ear recognition task, several metrics have been proposed, which will be presented here in the context of their dependence on performance level, \( P_0 \), where \( P_0 = (\%_{\text{RE}} + \%_{\text{LE}})/2 \). (1) \( d \) is the simple difference score, \( \%_{\text{RE}} - \%_{\text{LE}} \), and \( d \) is unconstrained by performance level only when \( P_0 = 50\% \). (2) Because performance level imposes a ceiling on \( d \), Halwes (1969) proposed an “index of laterality,” which will be symbolized as \( L_i \), \( L_i = [\text{(R - L)/(R + L)]100} \), and the analysis is confined to those trials in which only one response is correct. (3) \( POC \) reflects the percentage of correct responses (Harshman and Krashen, 1972). (4) \( POE \) reflects the percentage of erroneous responses (Harshman and Krashen, 1972). (5–6) \( POE' \) and \( POE'' \) are linear transformations of \( POE \) and \( POE' \) (Repp, 1977). \( POE' \) is unconstrained when \( P_0 < 50\% \), and \( POE'' \) is unconstrained only when \( P_0 = 50\% \). (7) \( \phi \) is the geometric mean of \( POE' \) and \( POE'' \) (Kuhn, 1973) and is unconstrained only when \( P_0 = 50\% \). (8) \( e \) is a disjunctive use of \( POE' \) and \( POE'' \), meaning that \( e = POE' \) when \( P_0 < 50\% \) and \( e = POE'' \) when \( P_0 > 50\% \). Thus, \( e \) is independent of \( P_0 \).

The maximum value of the various indices can be constrained by \( P_0 \), but the dependence apparently is not particularly strong. Speaks, Niccum, and Carney (1982) reported that the correlations of the various metrics with \( P_0 \) ranged from \(-0.18 \) to \(+0.17 \). Moreover, all inter-correlations among the metrics were \( > 0.95 \).

What, then, is the metric of choice? We (Speaks, Niccum, and Carney, 1982) have argued that the ideal metric should reflect two properties. One is the proportion of all trials on which an ear advantage occurs, which is the ratio of single-correct (SC) trials to total trials: \( p(SC) = SC/(SC + DC + DE) \), where DC refers to double-correct trials and DE refers to double-error trials. The second property is the magnitude of ear advantage, \( p(EA_{\text{mag}}) \), on those trials in which an ear advantage occurs, which is the ratio of the difference between single-correct trials for the right ear and single-correct trials for the left ear to total single-correct trials: \( p(EA_{\text{mag}}) = (SC_{\text{RE}} - SC_{\text{LE}})/SC \), and \( p(EA_{\text{mag}}) = L_i \). If we assume that the two properties are of equal importance, we can derive a single metric for the ear advantage by computing the product of \( p(SC) \) and \( p(EA_{\text{mag}}) \), and that weighted value = \( d \), the simple ear-difference score. Finally, if we assume that scores for each ear and the ear advantage are normally distributed for the individual listener, the utility of \( d \) can be improved by a z-score transformation, a \( d' \)-like statistic that incorporates variability of the ear advantage for the individual listener as an error term: \( d_Z = (\%_{\text{RE}} - \%_{\text{LE}})/\sigma \), where \( \sigma = ([\sigma^2_{\text{RE}} + \sigma^2_{\text{LE}}]/2)^{1/2} \).

**Reliability of Ear Advantage Scores, d**

The reliability of ear-advantage scores has been reported principally for the simple ear-difference score, \( d \), Ryan and McNeil (1974) tested listeners with two blocks of 30 CV nonsense syllables per block and reported a correlation between blocks of \( +0.80 \). That outcome compares favorably with the test-retest correlation of \( +0.74 \) reported by Blumstein, Goodglass, and Tartter (1975). These authors emphasized, however, that the correlation coefficient might not be sensitive to reversals in direction of the advantage between test and retest: an REA on one block and an LEA on a second block. In fact, reversals occurred for 29% of their listeners. Pizzamiglio, DePascalis, and Vignati (1974) reported a similar outcome: 30% of their listeners who were tested with digits evidenced a reversal in direction of the ear advantage. Other investigators have reported test-retest correlations of \( +0.64 \) (Catlin, Van Derveer, and Teicher, 1976) and \( +0.66 \) (Speaks and Niccum, 1977).

Repp (1977) contended that the poor reliability likely was due in part to an insufficient number of trials. He applied the Spearman-Brown predictive formula to the Blumstein, Goodglass, and Tartter (1975) data and estimated that a reliability coefficient of \( +0.90 \) should be realized with a 240-trial test (eight blocks of 30 pairs of syllables per block). Speaks, Niccum, and Carney (1982) provided empirical support for Repp’s prediction. They tested 24 listeners with 20 blocks of 30 pairs of CV syllables per block. The split-half reliability coefficient was only \(+0.62 \) when scores for block 1 were compared with scores for block 2. The coefficient improved to \(+0.91 \), however, when scores for blocks 1 through 6 were compared with scores for blocks 7 through 12. Moreover, the standard error of measurement diminished from 13.09 for the block 1–2 comparison to 5.66 for the block 1–6 and 7–12 comparison. They concluded that use of fewer than 180 listening trials is likely to generate unreliable data.

**Properties of the Ear Advantage**

An examination of scores of reports on the dichotic ear advantage with CV nonsense syllables shows one fairly consistent result. The mean ear advantage for a group of listeners is almost always an REA with a magnitude in the range of \( 5\%–12\% \). In most instances, however, there have been an insufficient number of listening trials to permit a more detailed analysis of relevant statistical properties. We will refer principally then to an experiment reported by Speaks, Niccum, and Carney (1982) on 24 listeners with the two-ear recognition task. Each listener received 20 blocks of 30 pairs of CV nonsense syllables per block (600 dichotic trials per listener). Thus, the properties of interest were derived from 1200 responses for each listener and from 28,800 responses (24 listeners \( \times \) 1200 per listener) for the group. We will note only a few of the more salient properties.

1. Mean percent scores for the group were 71.7% for the RE and 65.1% for the LE, and the ear advantage, \( d \),
was therefore +6.6% (an REA). The RE and LE scores for the 20 listening blocks presented to the 24 listeners were plotted as cumulative percentage distributions and fitted with normal integrals. This analysis suggested that there are two overlapping normal curves with different means (RE = 71.7%, LE = 65.1%) but equal variances (σ² RE = 8.9%, σ² LE = 8.8%). Thus, a true ear advantage equals the difference between the means of the RE and LE probability density functions.

2. The block-to-block measures of the ear advantage also were distributed as a normal curve with a mean of 6.6% and a mean intralistener standard deviation of 10.8% ± 2.5%, where the mean sigma was calculated as the square root of the mean of the variances.

3. The mean/sigma ratio for the group of listeners was very small: 6.6%/10.7% = 0.62. Thus, the typical listener with a small ear advantage, on the order of 6.6%, should be expected to evidence block-to-block reversals in the direction of ear advantage such as those that have been reported in the literature. The relation between the size of the mean/sigma ratio and reversals in ear advantage can be illustrated by results obtained for three of the 24 listeners. For L-1, d = 23.3% (REA), σ = 9.7%, and X/σ = 2.40. An REA occurred on all 20 listening blocks. For L-2, d = −26.5% (LEA), σ = 10.1%, and X/σ = −2.62. An LEA occurred on 19 of the 20 listening blocks, and a NoEA was observed on one block. For L-3, d = 0.5% (a nonsignificant REA), σ = 10.7%, and X/σ = 0.05. Of the 20 listening blocks, 4 were LEA, 10 were NoEA, and 6 were REA.

How Should the Dichotic Ear Advantage Be Interpreted?

An REA is widely accepted as evidence of left hemispheric dominance for processing, and an LEA is viewed as evidence of right hemispheric dominance for processing. The REA commonly reported for speech is thought to arise from “loss of information” during interhemispheric transmission. The REA is a widely accepted as evidence of left hemispheric dominance for processing, and an LEA is viewed as evidence of right hemispheric dominance for processing. The REA commonly reported for speech is thought to arise from “loss of information” during interhemispheric transmission. The REA is thought to be evidence of right hemispheric dominance for processing.

There is general agreement that about 95% of the right-handed population (and perhaps 70% of the left-handed population) is left hemispheric dominant for the processing of speech and language (Penfield and Roberts, 1959; Geschwind and Levitsky, 1968; Annett, 1972). If the direction of ear advantage reflects the direction of hemispheric laterality, it seems reasonable to assume that something approaching 95% of the right-handed population should have an REA for speech signals presented dichotically. No outcome approaching 95% has been reported, and unfortunately, the significance of the observed ear advantage for individual listeners has rarely been tested statistically. In the Speaks, Niccum, and Carney (1982) experiment in which the two-ear recognition task was used with 24 listeners, 18 had an observed REA, but the advantage was only significant (McNemar’s χ² for correlated proportions) at the 0.05 level or less for 12 listeners (50%). Three (12.5%) of the 24 listeners had a significant LEA. Katsuki et al. (1984) tested 20 listeners with a yes/no target-monitoring task and the ear advantage for individual listeners was tested on transformations (Φ = [2 arcsin P(C)max]1/2) on P(C)max for each ear. Thirteen (65%) of 20 listeners had a significant REA, and two (10%) had a significant LEA.

A similar outcome was reported by Wexler, Halwes, and Heninger (1981). They tested 31 listeners and found that 14 (45%) had a significant REA and one (3%) had a significant LEA. They, however, placed a different interpretation on their data by disregarding the fact that 16 listeners did not have a significant ear advantage and emphasizing that 14 (93%) of the 15 listeners who had a significant ear advantage had a significant REA. To justify dismissing the 16 listeners with no ear advantage from their analysis assumes that processing for speech must always be lateralized to either the left or the right hemisphere and that failure to obtain an observed ear advantage must be due principally to measurement error. They suggested that increasing the length of the dichotic test might reduce measurement error and lead to a larger number of listeners having a significant REA. But, as we have seen, even with 20 listening blocks (Speaks, Niccum, and Carney, 1982), the mean ear advantage for a group of 24 listeners was only 6.6% (REA), and only 12 of 24 listeners had a significant REA. We acknowledge that most, perhaps 80% or so, listeners who have a dichotic ear advantage have an REA rather than an LEA. From a different perspective, however, no more than two-thirds of listeners tested appear to have an REA. Because a large proportion of the right-handed population is known to be left hemispheric dominant for speech and language processing, but a much smaller proportion evidence an REA, we believe that speculations about the neurological bases for listener responses to dichotic stimulation are at best fragile.

—Charles Speaks

References


Electrocochleography

Electrocochleography (ECochG) refers to the general method of recording the stimulus-related potentials of the cochlea and auditory nerve. The product of ECochG—the electrocochleogram, or ECochGm—is shown in Figure 1. As depicted in this figure, the components of interest may include the cochlear microphonic (CM), cochlear summating potential (SP), and auditory nerve action potential (AP). Detailed descriptions of these electrical events are abundant in the hearing science literature and are beyond the scope of this review. For a more thorough discussion of the history of these potentials as recorded in humans, see Ferraro (2000).

The popularity of ECochG as a clinical tool emerged in the early 1970s, following the discovery and application of the auditory brainstem response (ABR). The development and refinement of noninvasive recording

![Figure 1. Components of the click-evoked human electrocochleogram. Top tracings display responses to rarefaction (R) and condensation (C) polarity clicks. Adding separate R and C responses (middle tracing) enhances the cochlear Summating Potential (SP) and auditory nerve Action Potential (AP). Subtracting R and C responses (bottom tracing), enhances the Cochlear Microphonic (CM). (From American Speech-Language-Hearing Association, 1988, p. 9, based on data from Coats, 1981.)](Image 340x167 to 520x395)
techniques also has facilitated the current clinical use of ECochG.

The technical capability to record cochlear and auditory nerve potentials in humans has led to a variety of clinical applications for ECochG. Among the more popular applications are

1. To diagnose, assess, and monitor Ménière’s disease/endolymphatic hydrops (MD/ELH) and to assess and monitor treatment strategies for these disorders
2. To enhance wave I of the ABR
3. To monitor cochlear and auditory nerve function during operations that involve the auditory periphery (Ruth, Lambert, and Ferraro, 1988; Ferraro and Krishnan, 1997).

**ECochG Recording Techniques**

*Transtympanic Versus Extratympanic ECochG.* The terms transtympanic (TT) and extratympanic (ET) refer to the two general techniques for recording ECochG. The TT approach involves passing a needle electrode through the TM to rest on the cochlear promontory, whereas ET recordings are generally made from the surface of the ear canal or TM. The primary advantage of TT ECochG is that this “near-field” approach produces large components with relatively little signal averaging. The major limitation of TT ECochG is that it is invasive. ET recordings require more signal averaging, and the components tend to be smaller in magnitude than their TT counterparts. However, this approach is generally painless and can be performed by nonphysicians in non-medical settings.

TM offers a good compromise between ear canal and TT recording sites with respect to component magnitudes and signal averaging time without being invasive or painful (Lambert and Ruth, 1988; Ferraro, Blackwell, et al., 1994; Ferraro, Thedinger, et al., 1994). In addition, the waveform patterns that lead to the interpretation of the TT ECochGm are preserved in TM recordings (Ferraro, Thedinger, et al., 1994).

Figure 2 is a drawing of the TM electrode (or “tymptrode”) used in our clinic and laboratory, which is a modification of the device described several years ago by Stypulkowski and Staller (1987). Details regarding the fabrication and placement of the tymptrode can be found in Ferraro (1997, 2000).

**Recording Parameters.** ECochG components occur within the first few milliseconds of electrophysiological activity following stimulus onset. Table 1 illustrates the parameters used in our laboratory and clinic for recording the SP and AP together, which is usually the pattern of interest when ECochG is used in the diagnosis of MD/ELH. It is important to note that the bandpass setting of the analog filter of the preamplifier must be wide enough to accommodate the recording of both direct and alternating current potentials (i.e., the SP and AP, respectively).

| **Stimulus Considerations.** The broadband click tends to be the most popular stimulus for short-latency AEPs because it excites synchronous discharges from a large population of neurons to produce well-defined neural components. However, the brevity of the click makes it a less than ideal stimulus for studying cochlear potentials such as the CM and SP, whose durations are stimulus dependent. The use of tonal stimuli can overcome some of these limitations, and also provide for a higher degree of response frequency specificity than clicks (Durrant and Ferraro, 1991; Ferraro, Blackwell, et al., 1994; Ferraro, Thedinger, et al., 1994; Margolis et al., 1995). Stimulus polarity is an important factor for ECochG. Presenting clicks or tonebursts in alternating polarity inhibits the presence of stimulus artifact and CM, which are both dependent on stimulus phase. Thus, alternating-polarity stimuli may be preferable when the amplitudes of the SP and AP are of interest (as in the determination of the SP/AP amplitude ratio for the diagnosis of MD/ELH). |

**Table 1. ECochG Recording Parameters**

| **Electrode Array** | **Signal Averaging Settings** | **Stimuli** | **Recording Parameters.** ECochG components occur within the first few milliseconds of electrophysiological activity following stimulus onset. Table 1 illustrates the parameters used in our laboratory and clinic for recording the SP and AP together, which is usually the pattern of interest when ECochG is used in the diagnosis of MD/ELH. It is important to note that the bandpass setting of the analog filter of the preamplifier must be wide enough to accommodate the recording of both direct and alternating current potentials (i.e., the SP and AP, respectively). |

| **Electrode Array** | **Signal Averaging Settings** | **Stimuli** | **Recording Parameters.** ECochG components occur within the first few milliseconds of electrophysiological activity following stimulus onset. Table 1 illustrates the parameters used in our laboratory and clinic for recording the SP and AP together, which is usually the pattern of interest when ECochG is used in the diagnosis of MD/ELH. It is important to note that the bandpass setting of the analog filter of the preamplifier must be wide enough to accommodate the recording of both direct and alternating current potentials (i.e., the SP and AP, respectively). |

<table>
<thead>
<tr>
<th><strong>Primary (++)</strong></th>
<th><strong>Secondary (–)</strong></th>
<th><strong>Common</strong></th>
<th><strong>Timebase</strong></th>
<th><strong>Amplification Factor</strong></th>
<th><strong>Filter Bandpass</strong></th>
<th><strong>Repetitions</strong></th>
<th><strong>Type</strong></th>
<th><strong>Duration (BBC)</strong></th>
<th><strong>Envelope (TB)</strong></th>
<th><strong>Polarity</strong></th>
<th><strong>Repetition Rate</strong></th>
<th><strong>Level</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tympamic membrane</td>
<td>Contralateral earlobe or mastoid process</td>
<td>Nasion or ipsilateral earlobe</td>
<td>10 milliseconds</td>
<td>50,000–100,000 × (Extratympanic—ET)</td>
<td>5 Hz–3000 Hz</td>
<td>750–1000</td>
<td>Broadband Clicks (BBC), Tonebursts (TB)</td>
<td>100 microseconds electrical pulse</td>
<td>2 millisecond linear rise/fall, 5–10 msec plateau</td>
<td>Rarefaction and Condensation (BBC), Alternating (TB)</td>
<td>11.3/sec</td>
<td>95–85 dB HL (125–115 dB pe SPL)</td>
</tr>
</tbody>
</table>
ELH). Recording separate responses to condensation and rarefaction clicks also is useful, as certain subjects with MD/ELH display abnormal AP latency differences to clicks of opposing polarity (Margolis et al., 1992; Margolis et al., 1995; Orchik, Ge, and Shea, 1997; Sass, Densert, and Arlinger, 1997).

When ECochG is performed to help diagnose MD/ELH, stimulus presentation should begin at a level near the maximum output of the stimulus generator to evoke a well-defined SP-AP complex. Masking of the contralateral ear is not a concern for conventional ECochG since the magnitude of any electrophysiological response from the nontest ear is very small and ECochG components are generated prior to crossover of the auditory pathway.

**Interpretation of the ECochGm**

Figure 3 depicts a normal ECochGm to click stimuli recorded from the TM. Component amplitudes can be measured from peak to peak (left panel) or using a baseline reference (right panel). AP-N1 latency is measured from stimulus onset to the peak of N1 and should be identical to the latency of ABR wave I at the same stimulus level. When using a tubal insert transducer (highly recommended), these values will be delayed by a factor proportional to the length of the tubing. Although labeled in Figure 3, N2 has received little interest for ECochG applications.

Also as shown in Figure 3, SP and AP amplitudes are made from the leading edge of both components. The resultant values are used to derive the SP/AP amplitude ratio, which ranges from approximately 0.1 to 0.5 in normal subjects.

Figure 4 depicts a normal ECochGm to a 2000-Hz toneburst presented in alternating polarity at 80 dB HL. Action Potential (AP) and its first negative peak (N1) are seen at the onset of the response. Summating Potential (SP) persists as long as the stimulus. SP amplitude is measured at midpoint of response (point B), with reference to a baseline value (point A). Amplitude (microvolts)/time (milliseconds) scale at lower right. (From Ferraro, Blackwell, et al., 1994, p. 19.)

**Clinical Applications**

**MD/ELH.** As mentioned earlier, ECochG has emerged as one of the more powerful tools in the diagnosis, assessment, and monitoring of MD/ELH, primarily through the measurement of the SP and AP. Examples...
of this application are shown in Figures 6 (click-evoked ECoGms), and 7 (toneburst-evoked ECoGms). The upper tracings in Figure 6 were measured from the promontory (TT), whereas the lower waveforms represent TM recordings. The SP/AP amplitude ratios are enlarged under each condition. The rationale for this finding remains unclear, but may relate to the nature of the SP as a distortion product of transduction processes in the cochlea. In particular, ELH may augment this distortion and thus increase the amplitude of the SP. Enlarged SP/AP amplitude ratios also have been reported for perilymphatic fistulas (Ackley, Ferraro, and Arenberg, 1994), which suggests that the fluid pressure of the scala media may be the underlying feature to which ECoG is specific.

Figure 7 illustrates the difference between right and left SPs evoked by 2000-Hz tonebursts in a patient with MD/ELH on the right side. A pronounced, negative SP is seen on the affected side of this particular patient, whereas the unaffected side shows a normal pattern. The reported incidence of an enlarged SP and SP/AP amplitude ratio in the general Ménière’s population is approximately 60%–65% (Gibson, Moffat, and Ramsden, 1977; Coats, 1981; Kumagami, Nishida, and Masaaki, 1982). Testing patients when they are experiencing symptoms of MD/ELH has been shown to improve this percentage (Ferraro, Arenberg, and Hassanein, 1985). Other approaches to increasing the sensitivity of ECoG include measuring the AP-N1 latency difference between responses to condensation versus rarefaction clicks (Margolis et al., 1992, 1995), and measuring the respective “areas” of the SP and AP to derive the SP/AP area ratio (Ferraro and Tibbils, 1999).

Enhancement of Wave I. In a high percentage of hard-of-hearing subjects, including those with acoustic tumors, wave I of the ABR may be unrecordable in the presence of wave V (Hyde and Blair, 1981; Cashman and Rossman, 1983). This situation precludes the measurement of the I–V and I–III interwave intervals, which are key features of the ABR for neurodiagnostic applications. Under these and other less than optimal recording conditions, using an ECoG approach for recording the ABR has considerable utility (Ferraro and Ferguson, 1989). Figure 8 exemplifies this application in a patient with hearing loss. The top tracing represents the ABR recorded with surface electrodes at the vertex and ear lobes, and wave I is absent in the presence of wave V. When the TM is used as a recording site, however (bottom tracing), wave I is clearly present, permitting the measurement of the I–V interwave interval.

Intraoperative Monitoring. Intraoperative monitoring of inner ear and auditory nerve status during operations that involve the peripheral auditory system has emerged as an important application for ECoG. Such monitoring usually is done to help the surgeon avoid potential trauma to the ear and nerve in an effort to preserve hearing, to identify anatomic landmarks (such as the endolymphatic sac), or to help predict postoperative outcome (Lambert and Ruth, 1988; Gibson and Arenberg, 1991; Wazen, 1994). Figure 9 illustrates a series of ECoG responses recorded from a patient undergoing endolymphatic shunt decompression surgery for treatment of MD/ELH. A reduction in the SP/AP amplitude ratio was observed during the course of surgery,
and this patient reported improvement in symptoms postoperatively.

**Summary**

ECochG continues to be a useful clinical tool in the identification and evaluation of inner ear and auditory nerve disorders. Although this article has addressed the currently more popular clinical applications of ECochG, others will emerge as our knowledge of auditory physiology continues to improve, and the technical aspects of recording the electrical events associated with hearing become more sophisticated. Current areas in need of additional research include the standardization of recording and stimulus parameters, and studies designed to improve the sensitivity of ECochG to MD/ELH and other cochlear disorders.

—John A. Ferraro

**References**


**Further Readings**


Electronystagmography

Electronystagmography (ENG) refers to a battery of tests used to evaluate the vestibular system. The tests include (1) the Dix-Hallpike test, (2) ocular motor tests, (3) a search for positional and/or spontaneous nystagmus, (4) the caloric test, and (5) the failure of fixation suppression test. During ENG testing, eye movement is measured to determine the presence of peripheral vestibular nerve and/or end-organ vestibular or central nervous system (CNS) dysfunction. Two methods for recording eye movement are electro-oculography (EOG) and video-nystagmography (VNG). EOG is a recording of the corneoretinal potential with surface electrodes placed on the face, whereas VNG measures eye movement through the use of infrared video cameras mounted in goggles worn by the patient during testing.

Dix-Hallpike Test

The Dix-Hallpike maneuver is a provocative positioning test for benign paroxysmal positioning vertigo (BPPV) (Dix and Hallpike, 1952). The patient is seated upright with the head turned 45° toward the test ear and then quickly lowered into a supine position with the head hanging off the bed or table and still positioned at the 45° angle. The diagnosis of BPPV is based on characteristic clinical findings on the Dix-Hallpike test. These findings include (1) torsional nystagmus and vertigo that occur when the patient is placed in the provoking position, (2) a delay in the onset of vertigo and nystagmus, and (3) a duration of vertigo and nystagmus of less than 1 minute.

Ocular Motor Tests

The purpose of ocular motor tests is to test non-vestibular eye movements. These movements include the saccadic system, the smooth pursuit system, the optokinetic (OKP) system, and the gaze-holding mechanism. Abnormalities in the ocular motor systems generally localize CNS lesions; however, acute peripheral vestibular lesions may also cause abnormal findings.

The saccadic system rapidly changes the direction of the eye to acquire the image of an object of interest. Disorders of the saccadic system can include slowing of saccadic eye movements, impaired saccadic accuracy, and impaired reaction time. The preferred stimulus for testing the saccadic system is a random sequence paradigm presented with an array of light-emitting diodes (LEDs) controlled by a computer (Baloh et al., 1975). Measurement parameters on the saccade test include latency, velocity, and accuracy (Fig. 1).

The smooth pursuit system allows continuous, clear vision of objects moving within the visual environment. Patients with impaired smooth pursuit require frequent corrective saccades to keep up with the target, producing cogwheeling or saccadic pursuit responses. Most computerized ENG systems use a sinusoidal paradigm that offers precise control of frequency and amplitude of smooth pursuit testing. The most important measurement parameter of smooth pursuit is gain. Gain is calculated as the ratio of eye velocity to target velocity. Figure 2 shows gain and EOG recording during smooth pursuit for a normal subject.

The optokinetic system serves to hold images of the environment on the retina during sustained head rotation. There are two optokinetic stimuli: (1) partial field devices, which include the light bar and a small motorized drum, and (2) full-field devices (preferred), such as an optokinetic projector or large optokinetic drum, that fill at least 90% of the visual field. The measurement parameter of the OKP test is gain (ratio of eye velocity to field velocity). Gain should be at least 0.5 and symmetrical.

The gaze tests determine whether a patient has gaze-evoked nystagmus. Gaze-evoked nystagmus occurs when a leaky neural integrator causes the eyes to drift back toward the primary position, necessitating corrective saccades. Thus, the eyes cannot hold their position when looking at an eccentric target. Gaze-evoked nystagmus can be caused by drugs, cerebellar disease, brainstem lesions, and acute peripheral vestibular lesions. The stimulus for the gaze tests is a light bar target positioned 20° or 30° right, left, up, or down from the center.
Figure 1. Sample of ENG recordings of saccades produced with a random-sequence paradigm in a normal subject. Data points represent peak velocity, accuracy, and latency measurements for each rightward and leftward saccade.

Figure 2. Sample of ENG recording of smooth pursuit produced with a sinusoidal paradigm in a normal subject. Data points represent gain values for each rightward and leftward eye movement across target frequency.
Spontaneous Nystagmus

To determine the presence of spontaneous nystagmus, the subject is seated upright with eyes closed, the head positioned straight ahead (0°), and the subject engaged in mental alerting tasks to prevent suppression of nystagmus. Spontaneous nystagmus is typically horizontal jerk nystagmus and usually occurs due to a lesion to the peripheral vestibular system causing an imbalance in the tonic signals arriving at the oculomotor neurons. The imbalance produces a constant drift of the eyes in one direction, interrupted by fast components in the opposite direction. If the imbalance results from a peripheral vestibular lesion, then the pursuit system can cancel it. Thus, when the patient opens his or her eyes, the nystagmus disappears. The clinical finding of spontaneous nystagmus suggests an uncompensated peripheral lesion, typically on the side opposite the direction of nystagmus. The proof of the side of lesion, however, lies in the caloric test.

Positional Nystagmus

Positional nystagmus occurs when a subject is placed in the following static positions: sitting; supine; supine, head turned left; supine, head turned right; right lateral; left lateral; pre-irrigation position; and head-hanging straight, right, and left. The patient is asked to close his or her eyes to eliminate the effects of visual suppression and to perform mental alerting tasks to avoid central suppression.

Positional (and spontaneous) nystagmus is classified according to the direction of the fast phase. The direction of nystagmus can be fixed or changing. Direction-changing nystagmus can be geotropic (beating toward the earth) or ageotropic (beating away from the earth). Direction-changing nystagmus is an abnormal, non-localizing finding that is most often associated with peripheral vestibular disease (Lin et al., 1986). Positional nystagmus can also be direction-changing in a single head position. This clinical finding is rare and indicates a central pathology. There is evidence that both structural and metabolic factors can alter the specific gravity of the cupula in the semicircular canals and cause positional nystagmus (Honrubia, 2000). Positional nystagmus can also be caused by brainstem lesions, and most clinicians place the burden of proof on the ocular motor function tests. That is, if the ocular motor tests are within normal limits, then it is doubtful that a brainstem lesion exists.

The Caloric Test

The caloric test (Fitzgerald and Hallpike, 1942) is the most important test in the ENG test battery and the only clinical vestibular test that can lateralize a vestibular deficit. The caloric test stimulates the horizontal semicircular canal and involves measurement of
slow-component eye velocity. The key principle of the caloric test is the convection current created in the horizontal semicircular canal by changing the temperature of the endolymph. This convection current causes utriculo-petal endolymph flow and increased neural firing of the primary vestibular afferent nerve during warm water or air irrigation, so that nystagmus beats toward the ear that is stimulated. Irrigation with cool water or air produces utriculofugal flow and a decrease in neural firing of the primary vestibular afferent nerve, so that nystagmus beats away from the stimulated ear.

The fundamental assumption of the caloric test is that all four caloric irrigations of a given patient are equally strong. Variables that affect the strength of the caloric stimulus include stimulus temperature, stimulus duration, flow rate or volume, mental alerting procedures, size and shape of the external auditory canal, type of stimulus, patient age, and patient medication. The slow-component velocity (SCV) of the nystagmus has proved to be the most sensitive parameter of vestibular function and is currently the standard metric used for evaluating the caloric test (Henricksson, 1956). Because average SCVs vary widely in normal subjects, a ratio of right ear to left ear responses is used to analyze the caloric test results (Barber and Wright, 1973). Normal and symmetrical bithermal caloric nystagmus is shown in Figure 3. Unilateral weakness is a caloric asymmetry that results when one labyrinth is less sensitive to caloric irrigation than the other labyrinth (Fig. 4). The unilateral weakness is calculated by determining the amount by which the average SCVs provoked by right ear irrigation differ in intensity from those provoked by left ear irrigations. The formula for unilateral weakness is:

\[
UW(\%) = \frac{(RW + RC) - (LW + LC)}{RW + RC + LW + LC} \times 100
\]

where RW, RC, LW, and LC are the peak SCVs of the responses to right warm, right cool, left warm, and left cool irrigations, respectively. Most laboratories use an interear difference of greater than 20%–25% to determine if a unilateral weakness exists (Baloh and Honrubia, 1990; Jacobson, Newman, and Kartush, 1997). Unilateral weakness identifies a peripheral vestibular deficit on the weak side and is the most definitive finding on the ENG test for identifying and lateralizing a peripheral vestibular deficit.

A spontaneous nystagmus, such as in the case of an acute peripheral vestibular lesion, can shift the baseline of the caloric response. The caloric responses will be skewed toward the direction of the spontaneous nystagmus. The intensity difference between right-beating and left-beating caloric nystagmus is called directional preponderance (DP). The formula for directional preponderance is:

\[
DP(\%) = \frac{(RW + LC) - (RC + LW)}{RW + LC + RC + LW} \times 100
\]

where RW, RC, LW, and LC are the peak SCVs of the responses to right warm, right cool, left warm, and left cool irrigations, respectively.
cool irrigations, respectively. A directional preponderance of 30% or greater is abnormal (Balogh and Honrubia, 1990). Directional preponderance is considered a nonlocalizing finding and usually reflects the presence of spontaneous or positional nystagmus.

Other caloric abnormalities include failure of fixation suppression and bilateral weakness. Failure of fixation suppression indicates that a patient is unable to suppress visually the nystagmus by more than 50% of peak SCV, suggesting central vestibular involvement. Bilateral weakness occurs with weak or absent caloric responses in both ears. Bilateral weakness usually indicates a bilateral peripheral vestibular deficit, but it also may occur with CNS pathology.

—Faith W. Akin

References


Further Readings


Frequency Compression

Frequency compression (or frequency lowering) is a general term applied to attempts to lower the spectrum of the acoustic speech signal to better match the residual hearing of listeners with severe to profound high-frequency sensorineural impairment accompanied by better hearing at the low frequencies. This pattern of hearing loss is common to a number of different etiologies of hearing loss (including presbycusis, noise exposure, ototoxicity, and various genetic syndromes) and arises from greater damage to the basal region relative to the apical region of the cochlea (see noise-induced hearing loss; ototoxic medications; presbyacusis). The major effect of high-frequency hearing loss on speech reception is a degraded ability to perceive sounds whose spectral energy is dominated by high frequencies, in some cases extending to 10 kHz or beyond. Perceptual studies have documented the difficulties of listeners with high-frequency loss in the reception of high-frequency sounds (including plosive, fricative, and affricate consonants) and have demonstrated that this pattern of confusion is similar to that observed by normal-hearing listeners deprived of high-frequency cues through the use of low-pass filtering (Wang, Reed, and Bilger, 1978). Traditional hearing aids attempt to treat this pattern of hearing loss by delivering frequency-dependent amplification to overcome the loss at high frequencies. Such amplification, however, may not lead to improved performance and has even been shown to degrade the speech reception ability of some listeners with severe to profound high-frequency loss (Hogan and Turner, 1998).
The goal of frequency lowering is to recode the high-frequency components of speech into a lower frequency range that is matched to the residual capacity of a listener’s hearing. Frequency lowering has been accomplished through a variety of different techniques. These methods have arisen primarily from attempts at bandwidth reduction in the telecommunications industry, rather than being driven by the perceptual needs of hearing-impaired listeners. This article summarizes and updates the review of the literature on frequency lowering published by Braida et al. (1979). For each of seven different categories of signal processing, the major characteristics of each method are described and a brief summary of results obtained with it is provided.

**Slow Playback**

The earliest approach to frequency lowering was the playback of recorded speech at a slower speed than that used in recording. Each spectral component is scaled lower in frequency by a multiplicative factor equal to the slowdown factor. Although this method is not suitable for real-time applications (because of the inherent time dilation of the resulting waveform) and leads to severe alterations in temporal relations between speech sounds, it is nonetheless important to understand its effects because it is a component of many other frequency-lowering schemes. An important characteristic of this method is its preservation of the proportional relationship between spectral components, including the relation between the short-term spectral envelope and the fundamental frequency (F0) of voiced speech. A negative consequence of proportional lowering, however, is the shifting of F0 into an undesirably low frequency range (particularly for male voices). Results obtained in studies of the effects of slow playback on speech reception conducted with normal-hearing listeners (Tiffany and Bennett, 1961; Daniloﬀ, Shriner, and Zemlin, 1968) indicate that reductions in bandwidth up to roughly 25% produce only small losses in intelligibility, bandwidth reductions of 50% cause moderate losses in intelligibility, and bandwidth reductions of 66% or greater lead to severe loss in intelligibility. These studies have shown that the voices of females and children are more resistant to lowering than male voices (presumably because the fundamental and formant frequencies are higher for women than for men), that the effects of lowering are similar for the reception of consonants and vowels, and that performance with lowered speech materials improves with practice. In a study of slow playback in listeners with high-frequency sensorineural hearing loss, Bennett and Byers (1967) found beneﬁcial effects for modest degrees of frequency lowering (up to a 20% reduction in bandwidth) but that greater degrees of lowering led to a substantial reduction in performance.

**Time-Compressed Slow Playback**

A solution to the time dilation inherent to slow playback was introduced by techniques that compress speech in time (Fairbanks, Everitt, and Jaeger, 1954) prior to the application of slow playback. Time compression can be accomplished in different ways, including the elimination of a fixed duration of speech at a given rate of interruption or eliminating pitch periods from voiced speech. When the time-compression and slow-playback factors are chosen to be equal, the long-term duration of the speech signal can be preserved while at the same time frequencies are lowered proportionally. Fundamental frequency can be affected differently, depending on the particular characteristics of the time-compression scheme, including being lowered, remaining unchanged, or being severely distorted (see Braida et al., 1979). The intelligibility of speech processed by this technique for normal-hearing listeners is similar to that described above for slow playback; that is, bandwidth reduction by factors greater than 20% lead to severe decrease in performance (Daniloﬀ, Shriner, and Zemlin, 1968; Nagafuchi, 1976). Results of studies in hearing-impaired listeners (Mazor et al., 1977; Turner and Hurtig, 1999) indicate that improvements for time-compressed slow-playback speech compared to conventional linear or high-pass ampliﬁcation may be observed under certain conditions. Small beneﬁts, on the order of 10–20 percentage points, are most likely to be observed for small amounts of frequency lowering (bandwidth reduction factors in the range of 10%–30%), for female rather than male voices, and for individuals who receive little aid from conventional high-pass ampliﬁcation. A wearable aid that operates on the basic principles of time-compressed slow playback (the AVR Transonic device) has been evaluated in children with profound deafness (Davis-Penn and Ross, 1993; MacArdle et al., 2001) and in adults with high-frequency impairment (Parent, Chmiel, and Jerger, 1997; McDermott et al., 1999). A high degree of variability is observed across studies and across subjects within a given study, with substantial improvements noted for certain subjects and negligible effects or degradations for others.

**Frequency Shifting**

Another technique for frequency lowering employs heterodyne processing, which uses amplitude modulation to shift all frequency components in a given band downward by a fixed displacement. This process leads to the overlaying, or aliasing, of high-frequency and low-frequency components. Aliasing is generally avoided by the removal of low-frequency components through filtering before modulation. Systems that employ shifting of the entire spectrum have a number of disadvantages: although temporal and rhythmic patterns of speech remain normal, the harmonic relationships of voiced sounds are greatly altered, fundamental frequency is severely modified, and low-frequency components important to speech recognition are removed to prevent aliasing. Even mild degrees of frequency shifting (e.g., a 400-Hz shift for male voices) have been found to interfere substantially with speech reception ability (Raymond and Proud, 1962).
Frequency Transposition

When frequency shifting is restricted to the high-frequency components of speech (rather than to the entire speech spectrum), the process is referred to as frequency transposition. This approach has been incorporated into several different wearable or desktop aids (Johansson, 1966; Velmans, 1973) whose basic operation involves shifting speech frequencies in the region above 3 or 4 kHz into a lower frequency region, and adding these processed components to the original unprocessed speech signal. Generally, the most careful and controlled studies of frequency transposition indicate that benefits are quite modest. Transposition can render high-frequency speech cues audible to listeners with severe to profound high-frequency hearing loss (Rees and Velmans, 1993); however, these cues may interfere with information normally associated with the reception of low-frequency speech components (Ling, 1968). There is evidence to suggest that transposition aids may be more useful in training in speech production than in improving speech reception in deaf children (Ling, 1968).

Zero-Crossing-Rate Division

Another approach to frequency lowering lies in an attempt to reduce the zero-crossing rate of the speech signal. In these schemes, bands of speech are extracted by filtering and the filter outputs are converted to lower frequency sounds having reduced zero-crossing rates. Evaluations of a system in which processing was applied only to high-frequency components and inhibited during voiced speech (Guttman and Nelson, 1968) indicated no benefits for processed materials on a large-set word recognition task for normal-hearing listeners with simulated hearing loss. Use of this system as a speech-production aid for hearing-impaired children indicates that, following extensive training, the ability to produce selected high-frequency sounds was improved, while at the same time the ability to discriminate these same words auditorily showed no such improvements (Guttman, Levitt, and Bellefleur, 1970).

Vocoding

An important class of frequency-lowering systems for the hearing-impaired is based on the channel vocoder (Dudley, 1939), which was originally developed to achieve bandwidth reduction in telecommunications systems. Vocoding systems analyze speech into contiguous bandpass filters whose output envelopes are detected and low-pass-filtered for transmission. These signals are then used to control the amplitudes of corresponding channels. For frequency lowering, the set of synthesis filters correspond to lower frequencies than the associated analysis filters. Vocoding systems appear to have a number of advantages, including operation in real time and flexibility in terms of the choice of analysis and synthesis filters (which can allow for different degrees of lowering in different regions of the spectrum as well as for independent manipulation of F0 and the spectral envelope). The effect of degree of lowering in vocoder-based systems appears to be comparable to that described above for slow playback and time-compressed slow playback (Fu and Shannon, 1999). A number of studies conducted with vocoder-based lowering systems have demonstrated improved speech reception with training, both for normal-hearing (Takefuta and Swiggart, 1968; Posen, Reed, and Braida, 1993) and for hearing-impaired listeners (Ling and Druz, 1967; McDermott and Dean, 2000). When performance with vocoding systems is compared to baseline systems employing low-pass filtering to an equivalent bandwidth for normal listeners or conventional amplification for impaired listeners, however, the benefits of lowered speech appear to be quite modest. One possible reason for the lack of success of some of these systems (despite the apparent promise of this approach) may have been the failure to distinguish between voiced and unvoiced sounds. Systems in which processing is suppressed when the input signal is dominated by low-frequency energy (Posen, Reed, and Braida, 1993) lead to better performance (compared to systems with no inhibitions in processing for voiced sounds) based on their ability to enhance the reception of high-frequency sounds while not degrading the reception of low-frequency sounds.

Frequency Warping

A more recent approach to frequency lowering incorporates digital signal-processing techniques developed for correcting “helium speech.” The signal speech is segmented pitch synchronously, processed to achieve nonuniform spectral warping, dilated in time to achieve frequency lowering, and resynthesized with the original periodicity. Both the overall bandwidth reduction and the relative compression of high- and low-frequency components can be specified. These methods roughly extrapolate the variance associated with increased length of the vocal tract and include the following characteristics: they preserve the temporal and rhythmic properties of speech, they leave F0 of voiced sounds unaltered, they allow for independent manipulation of F0 and spectral envelope, and they compress the short-term spectrum in a continuous and monotonic fashion. Studies of speech reception with frequency-warped speech indicate that spectral transformations that lead to greater lowering of the high frequencies relative to the low frequencies are superior to those with linear lowering or with greater lowering of low relative to high frequencies (Allen, Strong, and Palmer, 1981; Reed et al., 1983). Improvements in the ability to identify frequency-warped speech with training have been noted for normal and hearing-impaired listeners (Reed et al., 1985). Improved ability to discriminate and identify high-frequency consonants has been demonstrated with such warping transformations compared to low-pass filtering for substantial reductions in bandwidth (up to a factor of 4 or 5). Overall performance, however, is similar for lowering and low-pass filtering, owing to reduced
performance for the lowering schemes on sounds with substantial low-frequency energy.

Summary and Conclusions

Attempts at frequency lowering through a variety of different methods have met with only limited success. Frequency lowering leads to a reduction in bandwidth of the original speech signal and to the creation of new speech codes which may sound unnatural to the untrained ear. Evidence from a number of different studies indicates that performance on frequency-lowered speech can improve with familiarization and training in the use of frequency-lowered speech. Many of these same studies, however, also indicate that even after extended practice, performance with the coded speech signals does not exceed that achieved with appropriate baseline conditions (e.g., speech filtered to an equivalent bandwidth in normal-hearing listeners or conventional amplification with appropriate frequency gain characteristics in hearing-impaired listeners). Although frequency-lowering techniques can lead to large improvements in the reception of high-frequency sounds, they may at the same time lead to detrimental effects on the reception of vowels and consonants whose spectral energy is concentrated at low frequencies. Because of the need to use the region of low-frequency residual hearing for recoding high-frequency sounds, the low-frequency components of speech may be altered as well through the overlaying of coded signals onto the original unprocessed speech or through wholesale lowering of the entire speech signal. In listeners with high-frequency impairment accompanied by good residual hearing in the low frequencies, benefits for frequency lowering have been observed for listeners with severe to profound high-frequency loss using mild degrees of lowering (no greater than 30% reduction in bandwidth). For children with profound deafness (whose residual low-frequency hearing may be quite limited), frequency lowering appears to be more effective as a speech production training aid for specific groups of phonemes rather than as a speech perception aid.

—Charlotte M. Reed and Louis D. Braida

References


Reed, C. M., Schultz, K. I., Braida, L. D., and Durlach, N. I. (1985). Discrimination and identification of frequency-lowered speech in listeners with high-frequency hearing im-
Functional Hearing Loss in Children

Functional hearing loss (FHL) is frequently forgotten or misdiagnosed in the pediatric population, despite the fact that it is well documented (Bowdler and Rogers, 1989). The diagnosis is often missed in children because of lack of awareness of its manifestations (Pracy et al., 1996), its incidence (Barr, 1963), and its multiple causes (Broad, 1980).

Functional hearing loss is one of several terms used to describe a hearing loss that cannot be ascribed to an organic cause (Aplin and Rowson, 1986). In FHL, actual audiometric thresholds are inconsistent with the voluntary thresholds of a patient. The term pseudohypacusis was coined by Carhart (1961) to describe a condition in which a person presents with a hearing loss not consistent with clinical or audiologic evaluation. Nonorganic hearing loss is used interchangeably with FHL and pseudohypacusis in that these terms do not comment on the intent, conscious or subconscious, of the patient being tested (Radkowski, Cleveland, and Friedman, 1998).

Terms such as psychogenic hearing loss and conversion deafness imply that the cause of the auditory disturbance is psychological, whereas malingering suggests the conscious and deliberate adoption or fabrication of a hearing loss for personal gain (Stark, 1966). FHL often appears as an overlay on an organic impairment. As such, the term functional overlay is used to describe an exaggeration of an existing hearing loss.

The incidence of FHL in children is not well documented. Valid estimates are lacking, owing to lack of consensus on the definition of FHL and the absence of a concerted effort to collect such data. Estimates range from 1% to 12%, but the validity of these figures is undermined by the lack of standard criteria for diagnosing FHL in children (Broad, 1980). Pracy et al. (1996) suggest that the incidence is higher than expected and cannot be compared with the reported incidence in adults. A number of studies that do report incidence suggest that FHL occurs more than twice as often in girls as in boys (Dixon and Newby, 1959; Brockman and Hoversten, 1960; Campanelli, 1963). In pediatric studies of FHL, the condition is typically diagnosed in adolescence (Berger, 1965; Radkowski, Cleveland, and Friedman, 1998).

The ability to describe a characteristic audiometric and behavioral profile of FHL in children is made problematic by its multiple manifestations and definitions. Aplin and Rowson (1990) described four...
subgroups of FHL: (1) cases involving emotional problems that typically preceded the audiologic examination; (2) cases in which mild anxiety or conscious mechanisms produce a hearing loss during an audiologic examination; (3) cases of malingering (the hearing loss is deliberately and consciously assumed for the purposes of financial gain); and (4) cases in which the audiologic discrepancies are artifactual as a result of lack of understanding or inattention during audiometric testing.

The hallmark of FHL in children is inconsistency in audiometric test results. The diagnosis of FHL in children is generally easier than in adults, because children are less able to consistently produce erroneous results on repeated testing (Pracy et al., 1996). A common presentation of possible FHL is the child who demonstrates no difficulty in conversational speech but has a voluntary pure-tone audiogram that suggests difficulty in speech recognition (Stark, 1966). Speech audiometric results that are better than the pure-tone results are another common indicator of FHL in children (Aplin and Rowson, 1990). Behavioral responses to speech audiometry typical of FHL in children include responding to only one syllable of a test word or to only one phoneme presented at a given intensity (Pracy et al., 1996). Hosoi, Tsuta, Murata, and Levitt (1999) reported several indicators of FHL in children seen during audiometric testing. These indicators included (1) poor test-retest reliability, (2) a saucer-shaped audiometric configuration, (3) the absence of a shadow curve with a severe to profound unilateral hearing loss, and (4) a discrepancy between the speech reception threshold and the pure-tone average.

Misunderstanding, confusion, or unfamiliarity with the test directions or procedures must be determined before further testing for FHL is initiated. After reinstruction, modification of speech audiometry is undertaken when FHL in a child is suspected. Pracy et al. (1996) described successful use of the Fournier technique (1956), in which the level of speech presentation is varied, thereby tricking the patient into responding at a level at which he or she had not previously responded. Other techniques using speech audiometry include use of an ascending threshold determination technique (Harris, 1958) and presentation of informal conversation at levels below the voluntary pure-tone thresholds.

Hosoi et al. (1999) proposed suggestion audiometry as a useful technique to detect FHL in children as well as to determine true hearing levels. In suggestion audiometry, standard audiometric procedures are preceded by an information session on the benefits of using a hearing aid. The child is shown a hearing aid, given information about it, and ultimately wears the hearing aid in the off position during testing. The hearing aid is without tubing or an earmold; consequently, the earphone used during testing is placed over the auricle with the hearing aid. Hosoi et al. found that 16 of 20 children diagnosed with FHL showed a significant change in hearing level following the suggestion technique.

The use of adult-oriented tests of pseudohypacusis has had limited success in children demonstrating FHL. The Stenger test (Chaiklin and Ventry, 1965), while easy to conduct, is appropriate only for unilateral hearing losses. The Doerfler-Stewart test (1946), Bekesy audiometry (Jerger and Herer, 1961), and the Lombard test (Black, 1951) require special equipment and are rarely used in adults in the twenty-first century.

Cross-check procedures such as otoacoustic emissions and acoustic reflex testing can be used to determine the reliability of pure-tone thresholds (Radkowski, Cleveland, and Friedman, 1998). However, for actual threshold determination, the auditory brainstem response (ABR) measurement has been shown to be effective. Yoshida, Noguchi, and Uemura (1989) performed ABR audiometry with 39 school-age children presenting with suspected FHL and found normal hearing in 65 ears of 35 patients. Although ABR audiometry is not a hearing test, Sanders and Lazenby (1983) suggest it can be a powerful tool in the identification and quantification of FHL.

After detecting FHL and determining true hearing levels, Hosoi, Murata, and Levitt (1999) suggest it is essential that information about the cause of the functional hearing loss be obtained. Reports on FHL in children cite lack of attention, deflection of attention, school difficulties, a history of abuse, and emotional problems as possible causes, among others.

Barr (1963) found that the extra attention paid to children because of their purported hearing difficulty had encouraged them, consciously or unconsciously, to feign hearing loss. Children with previous knowledge of ear problems may use lack of hearing and withdrawal from communication as a response to problems experienced at school (Aplin and Rowson, 1986). A history of trauma immediately preceding the complaint of hearing loss was reported in 18 patients studied by Radkowski, Cleveland, and Friedman (1998). Drake et al. (1995) found that FHL may be an indicator of child abuse.

A high suspicion in approaching children will not delay early intervention in cases of an organic hearing loss, but failure to recognize FHL can be costly and potentially hazardous to the pediatric patient (Radkowski, Cleveland, and Friedman, 1998). Unnecessary radiographic studies or exploratory surgical intervention, inappropriate amplification, and financial and psychosocial costs are among the possible outcomes of inappropriate diagnosis of FHL in children. Spraggs, Burton, and Graham (1994) reported on five adult patients who underwent assessment for cochlear implantation and were found to have nonorganic hearing loss. Similar findings have not been reported in the pediatric population, but with the proliferation of cochlear implants in children, accurate diagnosis of hearing loss is essential.

Once the diagnosis of FHL in a child has been established, a nonconfrontational approach is recommended. Giving the child the opportunity to “save face” often can be achieved with retraining, reassuring, retesting, and convincing the child that the hearing loss will improve. Retesting with accurate results is often accomplished during one visit but may require follow-up visits.

Labeling the child a malingeringer is detrimental to the child and decreases the probability of obtaining accurate
behavioral thresholds. If the child is labeled as exhibiting FHL, the chance to back out gracefully is lost, and the functional component of the hearing loss may be solidified (Pracy et al., 1996). Pracy and Bowdler (1996) advocate an approach that treats the child as if the hearing loss were real, followed by the use of speech audiometry techniques to determine actual hearing thresholds. In recalcitrant cases, ABR may be required for threshold determination. Psychiatric referral is rarely necessary or desirable and should be reserved only for the intractable child (Bowdler and Rogers, 1989).

—Patricia McCarthy

References


Further Readings


Genetics and Craniofacial Anomalies

Genetic factors contribute to more than half of all congenital hearing losses and are also responsible for later-onset hearing losses. Understanding the factors underlying hereditary hearing loss requires locating the genes responsible for hearing loss and defining the specific mechanisms and functions of those genes. From a clinical standpoint, this information may contribute to improved management strategies for individuals with hereditary hearing loss and their families. Accurate determination of the auditory characteristics associated with various genetic abnormalities requires the use of measures sensitive to subtle aspects of auditory function.

Hereditary Hearing Loss. Congenital (hereditary) hearing loss occurs in approximately 1–2 of 1000 births, and at least 50% of all cases of hearing loss have a genetic origin (Morton, 1991). Although hereditary hearing losses may occur in conjunction with other disorders as part of a syndrome, the majority of cases are non-syndromic. Later-onset hereditary hearing loss occurs at various ages from the first decade to later in life.

Chromosomes. Human cells contain 23 pairs of chromosomes (22 pairs of autosomes and two sex chromosomes). Hereditary material in the form of DNA is carried as genes on chromosomes. Cells reproduce by mitosis (meiosis for the sex chromosomes), where chromosomes divide, resulting in two genetically similar
cells. Errors can occur during mitotic or meiotic division, resulting in cells with chromosomal abnormalities and an individual with a chromosomal defect.

**Genotype and Phenotype.** Genotype describes an individual’s genetic constitution. Phenotype relates to the physical characteristics of an individual and can include information obtained from physiological, morphological, and biochemical studies. Auditory tests contribute to the phenotypic description.

**Inheritance Patterns**

Hereditary hearing loss follows several patterns of inheritance. Autosomal recessive inheritance occurs in 70%–80% of individuals with nonsyndromic hearing loss. To display a recessive trait, a person must acquire one abnormal gene for the trait from each parent. Parents are heterozygous for the trait since they each carry one abnormal gene and one normal gene. Thus, recessively inherited defects appear among the offspring of phenotypically normal parents who are both carriers of a single recessive gene for the trait. When both parents are carriers, the chance of a child receiving two copies of the abnormal gene and displaying the phenotype is 25%. The parents’ chance of having a carrier child is 50%, and there is a 25% chance of having a child with no gene for the defect. In cases of nonsyndromic recessive hearing loss, a genetic source may be suspected in families with two or more occurrences of the disorder. Recessive inheritance occurs more commonly in nonsyndromic than in syndromic hearing loss.

In autosomal dominant inheritance, a single copy of an abnormal gene can result in hearing loss; thus, an affected parent has a 50% chance of passing that gene to their child. Autosomal dominant hereditary hearing loss occurs in approximately 15%–20% of nonsyndromic hearing loss and is more commonly associated with syndromic hearing loss. Other inheritance patterns that can result in hearing loss are X-linked, at a rate of 2%–3%, and mitochondrial, which occurs in less than 1% of cases.

**Variability in Hereditary Hearing Loss.** Phenotypic and genetic heterogeneity is pronounced, with reports of more than 400 forms of syndromic and nonsyndromic hereditary hearing loss (Gorlin, Toriello, and Cohen, 1955). Considerable variation exists among hereditary hearing losses, between dominant and recessive hearing losses, among various forms of either recessive or dominant hearing loss, and even among persons with the same genetic mutations. Furthermore, the same genes have been found responsible for both syndromic and nonsyndromic hearing loss, and have been associated with both autosomal dominant and recessive transmission.

Hereditary hearing losses range from mild to profound (Nance and Sweeney, 1975). In subjects with autosomal recessive nonsyndromic hereditary hearing loss, onset of the hearing loss tends to be congenital, severe to profound in degree, stable over time, and affecting all frequencies (Liu and Xu, 1994). Autosomal dominant nonsyndromic hereditary hearing loss tends to be less severe, more often delayed in onset, progressive, and affecting high frequencies. Patients with X-linked hearing loss generally have prelingual onset but are clinically diverse.

Mutations in the GJB2 (connexin 26) gene may explain greater than 50% of autosomal recessive deafness in some populations (Zelante et al., 1997). The GJB2 gene encodes the protein connexin 26 (Cx26), thought to be essential for maintenance of high potassium in the scala media of the inner ear. Several mutations in the GJB2 gene have been associated with hearing loss, and mutation sites vary among world populations. A 35delG mutation is common in some Mediterranean-based populations (Denoyelle et al., 1997), while a 167delT mutation is most common in the Ashkenazi Jewish population (Morell et al., 1998). Cx26 mutations are generally responsible for recessive deafness, although they have been observed in dominant deafness.

Hearing losses associated with Cx26 mutations are cochlear in nature but vary widely in degree, ranging from mild to profound, and stability (e.g., Cohn et al., 1999; Denoyelle et al., 1999; Mueller et al., 1999; Wilcox et al., 2000). Hearing losses resulting from the same genetic mutations show wide variability in degree and progression. Furthermore, audiometric characteristics are not directly linked to a particular type of mutation (e.g., Cohn et al., 1999; Sobe et al., 2000).

**Chromosomal Defects.** Down syndrome is the most common autosomal defect. The affected individual has an additional chromosome 21 (trisomy 21), for a total of 47 chromosomes, or a translocation trisomy. Down syndrome is characterized by mental retardation and a number of craniofacial and other characteristics. Hearing loss may be congenital, sensory, and there is a high incidence of middle ear disorders. Trisomy 13 and 18 syndromes, less common and with more dramatic abnormalities, are characterized by inner ear dysplasias involving the organ of Corti and stria vascularis, external and middle ear malformations, cleft lip and palate, and other defects.

** Syndromes Associated with Hearing Loss**

There are far too many syndromes associated with hearing loss to include in this brief entry. A useful method of classifying hearing loss was provided by Konigsmark and Gorlin (1976), where genetic and metabolic hearing losses were divided into major categories depending on the organ system or metabolic defect involved.

Usher syndrome is the most common syndrome associated with hearing loss and eye abnormalities, specifically retinitis pigmentosa. There are several types and subtypes and various genetic loci. Other syndromes involving vision are Cockayne syndrome and Alström disease, associated with retinal disorders. Treacher Collins syndrome, Goldenhar syndrome (hemifacial micro-
somnia), Crouzon syndrome (craniofacial dysostosis), Apert syndrome, otopalatal-digital syndrome, and osteogenesis imperfecta are all associated with musculoskeletal disease.

Waardenburg syndrome, characterized by displaced medial canthi, white forelock, heterochromia, and broad nasal root, is the most prominent hearing syndrome involving the integumentary system. Alport syndrome is a combination of progressive hearing loss, progressive renal disease, and ocular lens abnormalities. Pendred syndrome, mucopolysaccharidosis, and Jervell and Lange-Nielsen syndrome are associated with metabolic and other abnormalities. The diverse syndromes associated with neurological disorders include Friedreich ataxia and acoustic neuromas and neural deafness.

Craniofacial anomalies associated with hearing loss may be sporadic, inherited, due to disturbances during embryonic development, of toxic origin, or related to chromosomal abnormalities. These maldevelopments may be of unknown origin or related to known syndromes.

Inner ear dysplasias include Michel deafness, which is rare and involves complete inner ear dysplasia, Mondini deafness, and Scheibe deafness (Schuknecht, 1974). In Mondini deafness, the bony cochlear capsule is flattened, with underdevelopment of the apical turn of the cochlea and possible saccular and endolymphatic involvement. Hearing loss is typically moderate to profound but varies widely. Scheibe deafness involves the membranous portion of the cochlea and saccule, greater in basal portions, and is the most common of the inner ear dysplasias.

External and middle ear anomalies are associated with improper development of the first and second branchial clefts and arches, which are also responsible for lower jaw and other structures. Middle ear abnormalities include absence or fusion of the ossicles or abnormalities of the eustachian tube or middle ear cavity. Middle ear anomalies may be suspected whenever other branchial arch anomalies such as external ear atresia, cleft palate, micrognathia, Treacher Collins syndrome, Pierre Robin syndrome, and low-set auricles are present. Skeletal defects, such as those associated with Apert syndrome, Klippel-Feil syndrome, and Paget disease, and connective tissue disorders, such as those related to Hunter-Hurler or Möbius syndromes, may also indicate the presence of middle ear anomalies. Maldevelopment of the external ear includes preauricular tags, microtic or deformed pinna, or partial or complete atresia of the external canal. The presence of external or middle ear anomalies may indicate additional malformations or reduced hearing, depending on the structures and degree of involvement.

Evaluation and Classification of Hearing Loss

The characteristics of a hearing loss are important in understanding relationships, or lack of relationships, between genotype and phenotype. The audiogram provides a general description of the degree, configuration, frequency range, type, and progression of a hearing loss, and whether one or both ears are affected. Other, more sensitive measures (such as otoacoustic emissions, efferent reflexes, and auditory-evoked potentials) are necessary to understand the nature of a hearing loss in more detail.

The majority of hereditary hearing losses are non-syndromic, with no associated disorders that might raise the index of suspicion or aid in diagnosis. Furthermore, since the majority of nonsyndromic hearing losses are recessively inherited, parents are not affected by hearing loss. There may be no history of hearing loss in the family, so this risk factor would not be an indicator to raise suspicion of hearing loss. Thus, identification of nonsyndromic, and particularly recessively inherited, hearing loss is particularly challenging clinically.

See also SPEECH DISORDERS: GENETIC TRANSMISSION.

—Linda J. Hood

References


Audibility measures usually combine information about the availability of amplified acoustical speech cues with weighting factors proportional to the importance of those cues for speech recognition. The availability of cues may be limited by sensitivity thresholds or competing noises, as well as by the speech level. The importance of cues for speech intelligibility varies with frequency and type of speech. Studies of normal-hearing listeners and listeners with mild to moderate hearing impairments have shown that measures of weighted audibility can provide rather accurate predictions of speech intelligibility scores for these individuals in a laboratory setting. There are several well-researched approaches to obtaining audibility measures (e.g., Studebaker, 1992; ANSI, 1997). Some methods incorporate the effects of age, speech level, or hearing loss to improve the accuracy of speech intelligibility predictions for elderly hearing-impaired listeners who are exposed to high-level amplified speech.

Measuring the recognition of amplified speech is perhaps the most venerable approach to evaluating hearing aid fitting outcomes. Many standardized tests are available, in both open-set and closed-set varieties, with stimuli ranging from nonsense syllables to sentences. The popularity of speech intelligibility testing as a measure of outcome is rooted in its high level of face validity. The most frequently cited reason for obtaining amplification is a need to improve communication ability. A measure of improved speech understanding consequent on hearing aid fitting addresses that need in an attractively direct manner. For many years, this measure was the bedrock of hearing aid fitting outcome evaluation. Unfortunately, in order to achieve a useful level of statistical power, speech intelligibility tests must include a large number of test items. This requirement has limited the recent use of speech intelligibility testing mostly to research applications (e.g., Gatehouse, 1998).

The importance of producing amplified sounds that are acceptably loud has long been recognized. Although using loudness data to facilitate fitting protocols has been advocated for many years, interest in measuring the loudness of amplified sounds following the fitting has burgeoned since the widespread acceptance of wide dynamic range compression devices. With these instruments, it is appealing to assess the extent to which environmental sounds, including speech, have been “normalized” by amplification. Interest in this type of outcome data is increasing, and some of the measurement issues have been addressed (e.g., Cox and Gray, 2001).

Formal ratings of aspects of amplified sounds, such as quality, clarity, distortion, and the like, have been used frequently in assessing technical merit in research applications but have not often been advocated for clinical use. These types of measures are commonly performed with listeners supplying a rating on a semantic differential scale, such as the 11-point version developed by Gabrielsson and Sjogren (1979). This approach has the advantage of permitting quantification of dimensions of amplified sounds that are not psychoacoustically accessible via other metrics.
**Evaluation of Real-World Impact**

The real-life effectiveness of a hearing aid is measured using subjective data provided by the hearing-impaired person or significant others. Numerous questionnaires have been developed and standardized specifically for the purpose of assessing hearing aid fitting outcomes, and many others have been conscripted to serve this application (Noble, 1998; Bentler and Kramer, 2000). In addition to standardized questionnaires, there is strong support for use of personalized instruments in which the patient identifies the items and thus creates a customized questionnaire (e.g., Dillon, James, and Ginis, 1997). Regardless of which type of inventory is used, there are at least seven different domains of subjective outcomes of hearing aid fitting that can be assessed. They include residual activity limitations, residual participation restrictions, impact on others, use, benefit, satisfaction, and quality-of-life changes. Most inventories do not assess all of these domains.

Residual activity limitations are the difficulties the hearing aid wearer continues to have in everyday tasks such as understanding speech, localizing sounds, and the like. The activity limitations experienced by a specific individual will depend on the demands of that person’s lifestyle. Residual participation restrictions are the unresolved problems or barriers the hearing aid wearer encounters to involvement in situations of daily life. This also differs with individuals but can include such circumstances as participation in church services, bridge clubs, and so on. ICF (2001) provides a full discussion of activity limitations and participation restrictions.

Hearing impairment often places a heavy burden on the family and friends as well as on the involved individual. In fact, encouragement (or compulsion) by significant others is sometimes the major factor prompting an individual to seek a hearing aid. The relief provided by amplification for the problems in the family constellation (i.e., the impact on others) is an important outcome dimension but one that has received relatively little attention to date.

A measure of benefit quantifies change in a hearing-related dimension of functioning as a result of using amplification. Benefit may be measured directly, in terms of degree of change (small versus large), or it may be computed by comparing aided and unaided performance on a particular dimension. Typical dimensions on which subjective benefit is measured are activity limitations and participation restrictions. Hearing-specific questionnaires are typically used to quantify hearing aid benefit.

Sometimes general, non-hearing-specific questionnaires are used to determine changes that result from hearing aid provision. These types of data tend to be interpreted as reflecting changes in general quality of life. A recent large-scale study found that hearing aid use was significantly associated with improvements in many aspects of life, including social life and mental health (Kochkin and Rogin, 2000). Despite the importance of these effects for individuals, functional health status measures that are often used to gauge quality of life tend not to be sensitive to the changes that result from hearing aid use (Bess, 2000).

It is not unusual to observe that an individual who reports substantial hearing aid benefit is nevertheless not satisfied with the device or does not use amplification very often. These observations suggest that daily use and hearing aid satisfaction are additional, distinct dimensions of real-world outcome that require separate assessment (e.g., Cox and Alexander, 1999).

**Relationship Between Technical Merit and Real-World Impact**

Numerous studies have shown that measures of technical merit are not strongly predictive of real-world outcomes of hearing aid fitting (e.g., Souza et al., 2000; Walden et al., 2000). Principal components analyses in studies using multiple outcome measures often show that the two types of measures tend to occupy separate factors (e.g., Humes, 1999). Many researchers feel that both types of data are essential for a full description of hearing aid fitting outcome.

*See also hearing aids: prescriptive fitting.*

—Robyn M. Cox

**References**


Dillon, H., James, A., and Ginis, J. (1997). The Client Oriented Scale of Improvement (COSI) and its relationship to several other measures of benefit and satisfaction provided by hearing aids. *Journal of the American Academy of Audiology, 8,* 27–43.


**Further Readings**


---

**Hearing Aids: Prescriptive Fitting**

Prescriptive procedures are used in hearing aid fittings to select an appropriate amplification characteristic based on measurements of the auditory system. The advantages of using a prescriptive procedure as opposed to an evaluative or other approach are (1) they can be used with all populations in a time-efficient way, (2) they help the clinician select a suitable parameter combination from among an almost unlimited number possible in modern hearing aids and settings, and (3) they can be verified. On the negative side, there is little interaction with the client when fitting hearing aids according to a prescriptive procedure, and any two people with the same type of loss may have different preferences for the loudness and the tone of sounds.

More than fifty prescriptive procedures for fitting hearing aids have been presented. The procedures vary with respect to the type of amplification characteristic that is prescribed, the type of data the procedure is based on, and the aim of the procedure.

The parameter most commonly prescribed in hearing aids is gain as a function of frequency. In linear devices, only one gain/frequency response is prescribed, which applies to all input levels that do not cause the hearing aid to saturate. If the hearing aid is nonlinear (contains compressor amplifiers), gain varies as a function of both frequency and input level. In that case, gain/frequency curves are prescribed for different input levels, or the static compression parameters are prescribed for selected frequencies. To avoid excessive loudness when listening to high-intensity input levels, the maximum output of the hearing device must also be prescribed.

Some procedures prescribe the amplification characteristic based on threshold levels only. Others use supratheraloudness judgments, such as the most comfortable level (MCL), the loudness discomfort level (LDL), or the entire loudness scale. Supporters of threshold-based procedures argue that loudness data are difficult to measure and unreliable, especially in children and special populations, and that preferred gain and maximum output can be adequately predicted from threshold levels. The argument for using individually measured loudness data is that the fitting will be more accurate because hearing aid users with the same audiogram can perceive loudness differently. Table 1 lists some of the most widely used prescription procedures developed to date. They are categorized according to which parameters are prescribed and the data used.

Most procedures for fitting linear devices share the general aim of amplifying speech presented at an average level to a comfortable level situated approximately halfway between threshold and LDL. The rationale is that such a response provides optimum speech understanding and comfortable listening in general situations. Despite this common rationale, the assumptions and underlying operational principles behind each procedure vary, producing very different formulas. The assump-
The audibility of all speech components is important for speech understanding (e.g., DSL; Seewald, Ross, and Spiro, 1985). Speech discrimination is highest when speech is presented at levels above MCL (e.g., MSU; Cox, 1988). Speech is best understood when speech bands at different frequencies have equal loudness (e.g., NAL-R; Byrne and Dillon, 1986; and CID; Skinner et al., 1982). For hearing aid users with mild to moderate losses, speech presented at average input levels is restored to the MCL when providing gain equal to about half the amount of threshold loss (e.g., Berger, Hagberg, and Rane, 1977; and POGO; McCandless and Lyregaard, 1983). Some of these procedures take the shape of the speech spectrum into consideration when prescribing gain at each frequency (DSL, MSU, CID, and NAL-R), whereas others introduce a reduction in the low-frequency gain to avoid upward spread of masking from low-frequency ambient noise (Berger, POGO). Either way, the net result is that, even for a flat hearing loss, less gain is prescribed in the low than in the high frequencies. The NAL-R procedure differs from the other linear procedures in two respects. First, the gain prescribed at any frequency is affected by the degree of loss at other frequencies. Second, it is the only procedure that is well supported by direct empirical data (e.g., Berger, Hagberg, and Rane, 1977; and POGO; McCandless and Lyregaard, 1983).

Some of these procedures take the shape of the speech spectrum into consideration when prescribing gain at each frequency (DSL, MSU, CID, and NAL-R), whereas others introduce a reduction in the low-frequency gain to avoid upward spread of masking from low-frequency ambient noise (Berger, POGO). Either way, the net result is that, even for a flat hearing loss, less gain is prescribed in the low than in the high frequencies. The NAL-R procedure differs from the other linear procedures in two respects. First, the gain prescribed at any frequency is affected by the degree of loss at other frequencies. Second, it is the only procedure that is well supported by direct empirical data (e.g., Byrne and Cotton, 1988).

One procedure for fitting nonlinear devices, NAL-NL1 (Dillon, 1999), follows the common rationale of procedures for fitting linear devices by aiming at maximizing speech intelligibility for any input level. To avoid amplifying all input levels to a most comfortable level, which probably would make the loudness of environmental sounds unacceptable, the rationale uses the constraint that for any input level, the overall loudness of speech must not exceed normal loudness. The NAL-R procedure differs from the other linear procedures in two respects. First, the gain prescribed at any frequency is affected by the degree of loss at other frequencies. Second, it is the only procedure that is well supported by direct empirical data (e.g., Byrne and Cotton, 1988).

Table 1. An Overview of Widely Used Prescriptive Hearing Aid Fitting Procedures

<table>
<thead>
<tr>
<th>Linear Amplification</th>
<th>Nonlinear Amplification</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berger</td>
<td>FIG6 (Figure 6)</td>
<td>POGO</td>
</tr>
<tr>
<td>POGO, POGO II (Prescription of</td>
<td>NAL-NL1 (National Acoustic Laboratories’</td>
<td>NAL-SSPL</td>
</tr>
<tr>
<td>Gain/Output)</td>
<td>nonlinear formula, version 1)</td>
<td></td>
</tr>
<tr>
<td>NAL-R, NAL-RP (National</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acoustic Laboratories)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use supra-threshold</td>
<td></td>
<td></td>
</tr>
<tr>
<td>loudness data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CID (Central Institute for the Deaf)</td>
<td>LGOB (Loudness Growth of Octave Bands)</td>
<td>CID*</td>
</tr>
<tr>
<td>DSL* (Desired Sensation Level)</td>
<td>IHAFF (Independent Hearing Aid Fitting Forum)</td>
<td>MSU*</td>
</tr>
<tr>
<td>MSU* (Memphis State University)</td>
<td>DSL[i/o]* (Desired Sensation Level, input/output function)</td>
<td>IHAFF</td>
</tr>
<tr>
<td>ScalAdapt (Adaptive fitting by category loudness scaling)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Supra-threshold loudness data may be measured or predicted from threshold levels.

Consequently, the goal of achieving equal loudness is progressively relaxed within the NAL-NL1 rule as hearing loss increases.

The rationale behind most nonlinear prescription procedures, however, is loudness normalization. Examples are LGOB (Humes et al., 1996), FIG6 (Killion and Fikret-Pasa, 1993), IHAFF (Cox, 1995), DSL[i/o] (Cornelisse, Seewald, and Jamieson, 1995), and ScalAdapt (Kiessling, Schubert, and Archut, 1996). The assumption behind this rationale is that "normal hearing" is best for speech understanding and for listening to environmental sounds. Loudness normalization is achieved by applying the gain needed to make narrow-band stimuli of any input level just as loud for the impaired ear as they are for normal ears. This rationale maintains the interfrequency variation of loudness that normally occurs for speech (Fig. 1). Consequently, loudness normalization is not consistent with the principles of equalizing loudness across frequency and deemphasizing loudness at those frequencies where loss is greatest, principles that have emerged from research into linear amplification.

![Figure 1](image-url)  
**Figure 1.** Graph illustrating loudness perception of speech when the interfrequency variation of intensity for speech is maintained (loudness normalization) and when the intensity of speech bands has been equalized to maximize speech intelligibility.
Because of the different assumptions and principles used by the various procedures, different procedures prescribe different amplification characteristics for the same type of hearing loss (Figs. 2–5). The differences are more pronounced for flat and reverse sloping loss than for the more common high-frequency sloping loss.

A recent evaluation of loudness normalization versus speech intelligibility maximization suggests that when the difference in prescription between the two rationales is substantial, hearing aid users prefer and perform better with the speech intelligibility maximization rationale (Keidser and Grant, 2001).

The gain/frequency curves may be prescribed according to the acoustic input the client is likely to experience. Simple variations applied to the amplification characteristic prescribed to compensate for the hearing loss have proved useful for compensating for defined changes in the acoustic input (Keidser, Dillon, and Byrne, 1996).

Such variations can be programmed into different memories in a multimemory hearing aid.

Some procedures also prescribe the maximum output of the hearing aid known as the saturation sound pressure level (SSPL). It is important to have the output level of the hearing aid correctly adjusted. If the SSPL is too high, the hearing aid can cause discomfort or damage to the hearing aid user. On the other hand, if the SSPL is too low, the hearing aid user may experience insufficient loudness and excessive saturation. Most procedures that prescribe SSPL aim at avoiding discomfort. In those cases the SSPL is set equal to or just below the hearing aid user’s discomfort level; examples are CID, MSU, POGO, and IHAFF. Only one procedure, NAL-SSPL, considers both the maximum output level and the minimum output level and prescribes a level halfway between these two extremes (Dillon and Storey, 1998). This is also the only procedure for prescribing the output level.
that has been experimentally evaluated. It was found to provide an SSPL that did not require fine-tuning for 80% of clients (Storey et al., 1998).

Many prescription procedures target a sensorineural loss of mild to moderate degree. Appropriate adjustments to the prescriptions may be needed if prescribing amplification for clients with a conductive component (Lybarger, 1963), and a severe to profound loss (POGO II: Schwartz, Lyregaard, and Lundt, 1988; NAL-RP: Byrne, Parkinson, and Newall, 1990). Some adjustments are also needed for clients who are fitted with one versus two hearing aids (Dillon, 2001).

When the hearing aid has been adjusted according to the prescriptive procedures, the setting can be verified against the prescribed target, either in a hearing aid test box or in the real ear. Verifying the prescriptive parameters in the real ear allows individual configurations of the ear canal and the acoustic coupling between ear and hearing aid to be taken into consideration. For some clients fine-tuning may be needed after the client has tried the hearing aid in everyday listening environments.

The most commonly used prescription procedures are readily available in an electronic format, either as specifically designed computer programs, in programs provided by hearing aid manufacturers for fitting their programmable devices, or in equipment for measuring real-ear gain.

See also HEARING AID FITTING: EVALUATION OF OUTCOMES; HEARING AIDS: SOUND QUALITY.

—Gitte Keidser and Harvey Dillon

References
Byrne, D., and Dillon, H. (1986). The National Acoustic Laboratories’ (NAL) new procedure for selecting the gain and

Figure 4. Audiogram and prescribed insertion gain curves for a person with a gently sloping high-frequency hearing loss. For IHAFF, the targets are calculated based on average threshold-dependent loudness data (Cox, personal communication). Note that the NAL-NL1 rule does not prescribe insertion gain at frequencies where it is doubtful that amplification will contribute to speech intelligibility.

Figure 5. Audiogram and prescribed insertion gain curves for a person with a steeply sloping high-frequency hearing loss. For IHAFF, the targets are calculated based on average threshold-dependent loudness data (Cox, personal communication). Note that the NAL-NL1 rule does not prescribe insertion gain at frequencies where it is doubtful that amplification will contribute to speech intelligibility.


Further Readings


Hearing Aids: Sound Quality

The electroacoustic characteristics of a hearing aid differ considerably from those of a high-fidelity communication system. The hearing aid has a relatively narrow bandwidth, and the frequency response of the hearing aid is almost never flat. Because hearing loss usually increases with increasing frequency, the frequency response of the hearing aid typically provides increased gain as a function of frequency in order to obtain audibility for the hearing aid user. While hearing-impaired listeners may require considerable gain in order to obtain audibility, the output of the hearing aid must be limited in order to prevent the amplified sound from becoming uncomfortably loud when input levels are high.

A number of studies have been carried out to determine the effect of manipulating the electroacoustic characteristics of a hearing aid on sound quality. Sound quality is a multidimensional construct. Gabrielson and Sjogren (1979) identified eight important “dimensions” related to the manipulation of the frequency response of loudspeakers, headphones, and hearing aids. These dimensions are clarity, fullness, brightness, shrillness, loudness, spaciousness, nearness, and disturbing sounds (distortion). The “overall impression of quality” consists of some weighting of these various dimensions.

The effect of manipulating the frequency response of the hearing aid has been assessed in many studies. The bandwidth, amount of low-frequency versus high-frequency gain, and the smoothness of the frequency response will all affect the sound quality of speech. Several studies have shown that the presence of low-frequency content is a strong determinant of good sound quality for hearing-impaired listeners listening at comfortable listening levels (Punch, 1978; Punch et al., 1980; Punch and Parker, 1981; Tecca and Goldstein, 1984; Punch and Beck, 1986). However, the bandwidth yielding better sound quality changes as a function of the input level to the hearing aid and the amount of amplification provided by the hearing aid. When hearing-impaired subjects listen at higher levels, a frequency response with less low-frequency amplification yields better sound quality (Tecca and Goldstein, 1984).

There is also good agreement among studies that too much high-frequency emphasis degrades sound quality. This type of amplification is characterized by descriptions of sound as shrill, harsh, and tinny (e.g., Gabrielson and Sjogren, 1979; Gabrielson, Schenkman, and Hagerman, 1988). Thompson and Lassman (1970), Neuman and Schwander (1987), and Leijon et al. (1991) found that the sound quality of a flat frequency response was preferred to a response with extreme high-frequency emphasis. The results of these studies point to the need for balance between the low- and high-frequency energy for good sound quality. Of course, the optimum balance for any person depends on the way that person’s hearing loss varies with frequency. It has also been realized that the frequency response requirements for good sound quality for music differ from those for speech. An extended low-frequency response is more important for music than for speech (e.g., Franks, 1982).

A smoother frequency response has better sound quality than a frequency response with peaks. Davis and Davidson (1996) found that hearing-impaired listeners preferred to listen to speech processed through a hearing aid with a moderate amount of damping that smoothed the large resonant peak in the frequency response of the hearing aid. This preference was true for both male and female voices in quiet and in noise. Smoothing of the frequency response resulted in judgments of greater brightness, clarity, distinctness, fullness, nearness, and openness. Similarly, van Buuren, Festen, and Houtgast (1996) investigated the effect of adding single and multiple peaks to a smooth frequency response. Hearing-impaired subjects rated pleasantness on a five-point rating scale. The smooth frequency response was rated as having better sound quality than any of the frequency responses with peaks. Based on the results of the study, the researchers recommended that peak-to-valley ratios in the frequency response of the hearing aid should not exceed 5 dB.

In spite of the general agreement among studies about the effect of frequency response on sound quality, Preminger and Van Tasell (1995) have shown intersubject differences with regard to their ratings of the various dimensions of sound quality as a function of the manipulation of frequency response. They suggested that measures of speech quality be used to select among alternative frequency responses yielding similar speech intelligibility (close to 100%).

Much of the research described above was carried out with linear hearing aids, with testing carried out at a single input level. However, many current hearing aids are nonlinear, which means that the frequency response characteristics change as a function of the input level. Full evaluation of sound quality would require testing with various signals at multiple input levels and determining optimum sound quality at each level.

The method of output limiting is another hearing aid parameter that has an important effect on sound quality. Output limiting is used in a hearing aid to prevent amplified sounds from becoming uncomfortably loud and to protect the ear from excessively loud sounds that...
might cause further damage to the hearing. Major
methods of output limiting include peak clipping, com-
pression limiting, or wide dynamic range compression.

Peak clipping causes the generation of signals in the
output signal that are not in the input signal. Harmonic
distortion (integer multiples of the input signal) and
intermodulation distortion (combinations of the har-
monics caused by sums and differences of the harmonic
distortion) are caused by peak clipping. The coherence
between the input signal and the output signal is pre-
dictive of sound quality. Moderate amounts of peak
clipping degrade the sound quality of speech in quiet and
in noise, and the sound quality of music (Fortune and
Preves, 1992; Palmer et al., 1995; Kozma-Spytek, Kate,
and Revoile, 1996). There is also an interaction between
the frequency response shaping of the hearing aid and
the clipping level. This interaction is subject-dependent
(Kozma-Spytek, Kates, and Revoile, 1996). Fortune
and Preves (1992) found specifically that hearing
aids having less distortion (higher coherence) were per-
ceived as having better clarity and brightness, as pro-
ducing less discomfort, and as yielding better overall
sound quality.

Compression is a nonlinear form of amplification in
which the gain of the amplifier is decreased as the input
to the amplifier increases. Compression amplification
may be used to limit output (compression limiting) or
to fit a wide range of signals into the listener’s dynamic
range (wide dynamic range compression). In general,
compression limiting preserves sound quality better than
peak clipping. Compression limiting does not generate
harmonic and intermodulation distortion and yields
higher coherence values. Hawkins and Naidoo (1993)
compared the effect of asymmetrical peak clipping and
compression limiting on sound quality and clarity of
speech in quiet, speech in noise, and music. For both
sound quality and clarity, and for all three types of
stimuli, compression limiting was preferred to peak clip-
ing under conditions in which the hearing aid input was
high enough to cause limiting.

For hearing aids utilizing wide dynamic range com-
pression, compression ratio, attack, and release time all
affect the sound quality of the processed signal. The ef-
effect of compression variables also depends on whether
the signal of interest occurs in quiet or in noise. Several
studies have shown that high compression ratios have a
negative effect on sound quality (Neuman et al., 1994,
1998; Boike and Souza, 2000). Neuman and colleagues
(1994, 1998) found that compression ratios higher than
3:1 significantly degraded the sound quality of a single-
band-compression hearing aid. Compression ratios that
did not significantly degrade sound quality in quiet,
degraded sound quality in noise. The sound quality of
linear amplification (no compression) was preferred
when background noise levels were high. Boike and
Souza (2000) also found that speech quality ratings
decreased with increasing compression ratio for speech
in noise. Compression ratio did not significantly degrade
quality for speech in quiet. Research to determine the
effect of compression on specific dimensions of sound
quality revealed that clarity, pleasantness, background
noise, loudness, and overall impression all showed neg-
avative effects of increasing compression ratio (Neuman
et al., 1998).

Release time also affects sound quality. If short re-
lease times are used, low-level noise is amplified in the
pauses between words. This amplification of low-level
noise has been found to have a negative effect on the
perceived sound quality. Neuman and colleagues (1998)
found that hearing-impaired listeners’ ratings of the
clarity, pleasantness, and overall quality of speech in
quiet and speech in noise all decreased as release time
was decreased from 1000 ms to 60 ms (single-band-
compression hearing aid). Ratings of the amount of
background noise increased as release time decreased.

It is clear that the electroacoustic characteristics of a
hearing aid have a significant effect on sound quality.
Sound quality has been recognized as an important fac-
tor in the acceptability of a hearing aid to the user, and
because of individual differences among listeners, it has
been suggested that sound quality should be considered
a factor in hearing aid fitting (e.g., Gabrielson and
Sjogren, 1979; Kuk and Tyler, 1990; Preminger and Van
Tasell, 1995; Lunner et al., 1997). Past research has
shown that characteristics of the listeners, characteristics
of the signal being amplified, and characteristics of the
amplification system all affect sound quality. Applica-
tion of sound quality judgments to the fitting of non-
linear and multimemory hearing aids should be helpful in
determining the appropriate settings for these devices
(e.g., Keidser, Dillon, and Byrne, 1995).

See also HEARING AID FITTING: EVALUATION OF OUT-
COMES; HEARING AIDS: PRESCRIPTIVE FITTING.

—Arlene C. Neuman

References
ratio on speech recognition and speech-quality ratings with
wide dynamic range compression amplification. Journal of
Speech Language and Hearing Research, 43, 456–468.
performance with damped and undamped hearing aids by
listeners with sensorineural hearing loss. Journal of Speech
and Hearing Research, 35, 191–201.
ration and aided loudness discomfort. Journal of Speech
and Hearing Research, 35, 175–185.
Franks, J. R. (1982). Judgments of hearing aid processed mu-
sic. Ear and Hearing, 3, 18–23.
The effects of different frequency responses on sound qual-
ity judgments and speech intelligibility. Journal of Speech
and Hearing Research, 31, 166–177.
sound quality and clarity with asymmetrical peak clipping
and output limiting compression. Journal of the American
Academy of Audiology, 4, 221–228.
multiple frequency response characteristics. Ear and Hear-
ing, 16, 575–586.


Further Readings


Hearing Loss and the Masking-Level Difference

The masking-level difference (MLD) (Hirsh, 1948) refers to a binaural paradigm in which masked signal detection is contrasted between conditions differing with respect to the availability of binaural differences cues. The most common MLD paradigm has two conditions. In the first, NoSo, both the masker and signal are presented in phase to the two ears. In this condition, the composite stimulus of signal plus noise contains no binaural difference cues. In the second condition, NoSπ, the masker is presented in phase to the two ears, but the signal is presented 180° out of phase at the two ears. In this condition, the composite stimulus of signal plus noise contains binaural difference cues of time and amplitude. There are many other MLD conditions, but an underlying similarity is that all involve at least one condition in which the addition of the signal results in a change in the distribution of interaural time differences, interaural amplitude differences, or both interaural time and amplitude differences. For a broadband masker and a 500-Hz signal frequency, the threshold for the NoSπ condition is approximately 15 dB better than that for the NoSo condition, reflecting the sensitivity of the auditory system to the small interaural differences that are introduced when the Sπ signal is presented in the No noise. The magnitude of the MLD is most robust at relatively low signal frequencies, but under specific circumstances, the MLD can be quite large at high frequencies (Madden and Pasanen, 1978). Whereas the anatomical stage of processing most critical for the MLD has its locus in the auditory brainstem, the MLD also hinges upon the fidelity of more peripheral auditory processing.

Neurological Disorders. Some of the most prominent applications of the MLD to clinical populations have

489
concerned patients with lesions affecting the auditory nerve and auditory brainstem. The rationale for using the MLD in such cases was based on the assumption that the critical stages of auditory processing underlying the MLD occur in the low or mid-brainstem. It was reasoned that lesions affecting the transmission of fine timing information within this region would be associated with reduced MLDs. The results from several audiological investigations have supported this assumption. For example, reduced MLDs have been reported in listeners with tumors of the auditory nerve and low brainstem, and in listeners with multiple sclerosis (Quaranta and Cervellera, 1974; Olsen and Noffsinger, 1976; Olsen, Noffsinger, and Carhart, 1976; Lynn et al., 1981). Poor binaural performance in such cases has been attributed to gross changes in the temporal discharge patterns in the peripheral auditory nervous system, due to either physical pressure on the nerve or, in the case of multiple sclerosis, demyelination of low brainstem neural tissue. In additional support of the idea that the MLD is determined by relatively peripheral auditory processes, several studies have indicated that the MLD is usually not reduced in listeners having specifically cortical auditory lesions (e.g., Bocca and Antonelli, 1976).

Binaural tests other than the MLD have also been used to probe for the existence of peripheral auditory neural disorder. For example, a test of interaural time discrimination termed phase response audiometry has been applied to patients having neural lesions in the auditory periphery (Nilsson and Liden, 1976; Almqvist, Almqvist, and Johnson, 1989). In general, such patients have been found to have a reduced ability to discriminate changes in interaural time differences.

**Hearing Dysfunction Related to Aging.** Presbyacusis refers not only to the cochlea-based losses of threshold sensitivity that typically accompany the normal aging process but also to possible auditory neural dysfunction that may coexist with (or exist independently of) cochlear loss. In general, results from studies of the MLD in the elderly indicate reduced MLDs with advancing listener age. MLDs are often reduced in presbycusis listeners, particularly when hearing loss is present at the frequencies of the test stimulus. Of greater interest is the fact that MLDs are sometimes reduced in elderly listeners (Fig. 1) even when the audiograms of the listeners do not indicate an age-related hearing loss (e.g., Grose, Poth, and Peters, 1994). Such findings are usually interpreted in terms of abnormal auditory neural processing in the aging auditory system. The nature of the underlying neural abnormality accounting for the reduced MLDs in elderly listeners is unknown. It is possible that such a dysfunction could make a significant contribution to the overall hearing disability associated with aging, as the MLD measures the kinds of auditory function that underlie, at least in part, our abilities to localize sound sources and to hear desired signals in noise backgrounds.

**Cochlear Hearing Loss.** As reviewed above, the MLD has potential relevance to site of lesion clinical audio-logical testing because of its sensitivity to neural auditory dysfunction. Unfortunately, the MLD is affected by a wide range of hearing pathologies, making the clinical specificity of this test relatively poor. For example, the MLD is often reduced in listeners with cochlear hearing loss (e.g., Olsen and Noffsinger, 1976; Hall, Tyler, and Fernandes, 1984; Jerger, Brown, and Smith, 1984). MLDs are particularly likely to be reduced in cases of asymmetrical cochlear hearing loss, but reduced MLDs are also quite common in cases of symmetrical hearing loss. Although reduced MLDs in cochlear hearing loss may sometimes be accounted for in terms of a relatively low sensation level of stimulation or in terms of stimulation asymmetry, in some studies MLDs in listeners with cochlear hearing loss (particularly Ménière’s disease) are reduced more than would be expected from stimulus level and asymmetry factors (Schoeny and Carhart, 1971; Staffel et al., 1990). Whereas such findings undermine the MLD as a test that can differentiate between cochlear and retrocochlear sites of lesion, they point to the potential of this test for understanding the effects of cochlear hearing loss on the processing of binaural information.

One general finding associated with cochlear hearing loss is variability of results among listeners. This clearly holds true for the MLD. It is not presently obvious what accounts for the variability in the size of the MLD across listeners with cochlear hearing loss. Some possible factors include the cause of the hearing loss, whether a particular case of hearing loss is associated with a substantial reduction in the number of nerve fibers contributing information for binaural analysis, and whether the cochlear disease state may affect the symmetry of the frequency or place encodings at the two cochleae. Although there is controversy on this point, it has been speculated that some forms of cochlear hearing impairment may be associated with a reduced ability of the auditory nerve to phase lock (Woolf, Ryan, and Bone,
This could contribute to reduced MLDs, and could also account for the poor interaural time discrimination found in many cochlear-impaired listeners. It is likely that several causes contribute to reduced MLDs among cochlear-impaired listeners. The challenge of determining the specific mechanisms underlying the results in particular patients remains.

Conductive Hearing Loss. In some listening conditions, conductive hearing loss can be considered in terms of a simple attenuation of sound. In this sense, performance in an ear with conductive hearing loss would be expected to be similar to that in a normal ear stimulated at lower level. The situation for some aspects of binaural hearing is more complicated. If the conductive loss is different in the two ears, the associated attenuation will be asymmetrical. This asymmetry could reduce the efficiency of binaural hearing. Colburn and Hausler (1980) also pointed out that another possible source of poor binaural hearing in conductive impairment is related to bone conduction. For sound presented via headphones, both the air conduction route of stimulation and the bone conduction route of stimulation are theoretically relevant. In normal-hearing listeners, the influence of bone-conducted sound on the stimuli reaching the cochlea is probably of no material consequence. In cases of conductive hearing loss, however, the bone-conducted sound could have a significant effect on the composite waveforms reaching the cochlea and could materially affect the distribution of interaural difference cues. It is therefore possible that MLDs could be substantially reduced because of this factor in cases of conductive hearing loss.

Studies of binaural hearing in listeners with conductive hearing impairment have found that binaural hearing is relatively poor in many subjects (Jonkees and van der Veer, 1957; Nordlund, 1964; Quaranta and Cervelleta, 1974; Hall and Derlacki, 1986). The factors of hearing asymmetry and the bone conduction route of stimulation (discussed above) probably contribute to this poor binaural hearing. However, it is likely that additional factors are involved. One relevant finding is that binaural hearing does not always return to normal immediately following middle ear surgery. In studies of adults having otosclerosis, binaural hearing has been found to remain abnormal up to 1–2 years following the restoration of a normal audiogram (Hall and Derlacki, 1986; Hall, Grose, and Pillsbury, 1990; Magliulo et al., 1990). As indicated in Figure 2, reduced MLDs have also been found in children with normal audiograms at the time of testing but with a history of hearing loss due to otitis media with effusion (Moore, Hutchings, and Meyer, 1991; Pillsbury, Grose, and Hall, 1991; Hall, Grose, and Pillsbury, 1995). It has been speculated that some of the difficulties in the binaural hearing of conductively impaired listeners may be related to a reduction in the efficiency of the neural processing of binaural difference cues (Hall, Grose, and Pillsbury, 1995).

The results of MLD studies indicate that binaural analysis is often negatively affected in most general types of auditory dysfunction. Whereas poor performance is

![Figure 2](image.png)

**Figure 2.** MLDs for children without (open circles) and with (filled triangles) a history of otitis media with effusion. The data were obtained sequentially, before and after the children with a history of otitis media received tympanostomy tubes. The data for the normal control group are repeated in the six panels. The region between the lines represented the 95% prediction interval for the normal group. Data are adapted from Hall et al. (1995).
due to abnormal neural function in cases of retrocochlear hearing disorders, the reasons for poor performance in cases of cochlear and conductive loss are often less clear, particularly when degree and symmetry of hearing loss are not sufficiently explanatory. Identification of the particular factors resulting in abnormal MLD results in particular individuals remains a challenge.

—Joseph W. Hall

References


Further Readings


Hearing Loss and Teratogenic Drugs or Chemicals

Most causes of hearing loss in newborn infants are hereditary and cannot be prevented. However, about 30% of cases of hearing loss in newborns have been linked to teratogenic factors, and many of these cases are preventable. Teratogens are factors capable of causing physical defects in the developing fetus or embryo and are typically grouped into four categories: infectious, chemical, physical, and maternal agents. During intrauterine life, the fetus is protected from many teratogens by the placenta, which serves as a filter to prevent the toxic substances from entering the fetus’ system. The placenta, however, is not a perfect filter and cannot prevent entry of all teratogens. Prenatal susceptibility to teratogens and the severity of the insult are quite variable. Four factors believed to contribute to this variability are dosage of the agent, the timing of the exposure, the susceptibility of the host, and interactions with exposure to other agents. This entry discusses teratogenic chemicals that contribute to hearing loss in the newborn. It should be kept in mind that much is still unknown about chemical teratogens and their ultimate impact on the developing auditory system.

Drugs

Of the prescription or over-the-counter medications, one of the best-studied groups known to have an adverse effect on the auditory system is the aminoglycosides. There is considerable information on the ototoxic effects of aminoglycosides in adults. Some aminoglycosides are thought to be toxic to both the auditory and vestibular systems. Aminoglycosides can cross the placenta, causing sensorineural hearing loss and labyrinthe damage in the fetus. This fetal ototoxicity can occur even in the absence of ototoxicity in the mother. Intrauterine ototoxicity has been reported for streptomycin, dihydrostreptomycin, kanamycin, and gentamicin. It is important to note that most studies of intrauterine ototoxicity have been retrospective and have not controlled for other factors, such as diuretic use, that could have acted synergistically.

Isotretinoin is a prescriptive retinoid (a kind of vitamin A derivative) used to treat persistent and severe cystic acne. Its teratogenic effects include abnormalities of the ophthalmologic, cardiovascular, vestibular, auditory, and central nervous systems. Specific auditory system abnormalities include enlargement of the saccule and utricle, shortening of the cochlea, and malformation of the external ear (Schuknecht, 1993; Westerman, Gilbert, and Schondel, 1994). Both conductive and sensorineural hearing losses have been reported in newborns as a result of maternal use of isotretinoin during gestation. Proper distribution of isotretinoin currently requires informed patient consent, a negative pregnancy test in the 2 weeks before treatment is initiated, and the use of contraceptives from 1 month before to 1 month after use of the drug (Dyer, Strasnick, and Jacobson, 1998).

Probably the best-known nonprescription drug with devastating effects on the unborn is thalidomide. This over-the-counter sedative was available for a short period of time in the late 1950s. Prescribed for morning sickness, when it was consumed during a susceptible period of fetal development, the implications for the baby were catastrophic. When thalidomide was ingested during the first trimester, when fetal limbs differentiate, babies were born with limb buds rather than fully formed limbs. A variety of ear anomalies were also noted, including atresia of the external auditory meatus, cochlear malformations, and absent acoustic and vestibular nerves (Dyer, Strasnick, and Jacobson, 1998). Thalidomide was withdrawn from the market but has subsequently been reintroduced in lesser dosage for use in a number of immunological and inflammatory disorders.

Although numerous other prescription and over-the-counter drugs are known to have ototoxic effects in adults, their ototoxic potential in the fetus has not yet been demonstrated. Those agents currently under suspicion but not proven to be ototoxic in utero include, among others, a variety of antibiotics, including some aminoglycosides and tetracyclines; anti-inflammatory agents; chloroquine, an antimalarial drug; chemotherapeutic drugs; and diuretics. For additional reviews of prescription and over-the-counter ototoxic medications, see Dyer, Strasnick, and Jacobson (1998) and Strasnick and Jacobson (1995).

The effects of recreational drug use by pregnant women have been difficult to document, for a number of reasons. First, the effects are multifactorial, meaning that one drug can interact with another, and many consumers are polydrug users. In addition to using more than one drug at a time, drug-using mothers may not be receiving proper prenatal health care, another factor contributing to premature delivery that can compromise the health of a newborn. Furthermore, mothers who consume drugs may not accurately or honestly report the range and degree of their drug use.
Ethyl alcohol is clearly the most widely abused drug in the United States. Fetal alcohol syndrome, first described in the early 1970s, is a pattern of anomalies resulting from maternal consumption of alcohol during pregnancy. The exact amount of maternal alcohol consumption required to cause fetal damage is uncertain, but the effects are believed to be related more to the amount consumed than to the timing of the consumption (Dyer, Strasnick, and Jacobson, 1998). Approximately 2 in 1000 newborns suffer from fetal alcohol syndrome (Strasnick and Jacobson, 1995). The syndrome consists of multiple congenital anomalies, including prenatal and postnatal growth retardation, craniofacial dysmorphology, developmental delay, and behavioral aberrations. The characteristic cranial features include microcephaly, narrow forehead, small nose and midface, a long, thin upper lip, and micrognathia.

The primary auditory concern with children diagnosed with fetal alcohol syndrome appears to be conductive hearing loss secondary to recurrent otitis media (Church and Gerkin, 1988). This may be the result of first and second pharyngeal arch malformations leading to eustachian tube dysfunction. Sensorineural hearing loss has also been reported at higher rates in children with fetal alcohol syndrome than in the general population (Gerber, Epstein, and Mencher, 1995; Church and Abel, 1998).

Although animal research has demonstrated ototoxic effects when fetuses are exposed to cocaine, similar findings have not been demonstrated in humans. The incidence of prenatal cocaine exposure has been reported to range from 11% to 14% of live births. Pregnant women metabolize cocaine slower than other people, making them more sensitive to small amounts of the drug. Metabolites of cocaine have been found in the urine of exposed infants for up to 10 days after birth. If the cocaine-exposed infant is born full term with normal or near normal birth weight, there appears to be no peripheral hearing loss, according to auditory brainstem response studies (Cone-Wesson and Wu, 1992). It does appear, however, that neurotoxic effects are present in cocaine-exposed infants whether they are full-term or low birth weight (Cone-Wesson and Spingarn, 1990; Salamy et al., 1990). The ingestion of cocaine by pregnant women results in vasoconstriction of blood vessels delivering nutrients and oxygen to the developing fetus, which can cause hypoxia. As such, infants with intrauterine exposure to cocaine are at risk for a variety of disordered neurobehaviors. It is still unclear, however, whether infants of cocaine-abusing mothers who are born prematurely and of low birth weight are at greater risk for neurobehavioral problems than premature, low-birth-weight infants who are not exposed to cocaine.

Babies whose mothers are addicted to heroine or methadone are also addicted at birth and will show signs of narcotic withdrawal. The child may be hyperirritable for several months and may continue to be hyperactive. No direct link between maternal heroine use and congenital hearing loss has been made.

Other Chemical Teratogens

Limited data are available on other possible chemical teratogens. At least one study has supported the thesis that maternal use of trimethadione, an antiseizure medication, can occasionally result in hearing deficits in the fetus (Jones, 1997). Although extensive examination has not been conducted to date, maternal opium and nicotine use have not been found to result in peripheral hearing loss in newborns.

See also: Ototoxic Medications.

—Anne Marie Tharpe

References


Further Readings


Hearing Loss Screening: The School-Age Child

Among school-age children in the United States, it is estimated that nearly 15% have abnormal hearing in one or both ears (Niskar et al., 1998). With newborn hearing screening now available in nearly every state, many sensorineural hearing losses are identified prior to school entry. Even so, comprehensive hearing screening of school-age children is important, for several reasons. First, it will be years before universal infant hearing screening is fully implemented. Second, late-onset sensorineural loss may occur in the weeks or months following newborn screening, especially in young children with complicated birth histories (Centers for Disease Control and Prevention, 1997). Third, mild sensorineural loss can escape detection even when newborn hearing screening is provided (Joint Committee on Infant Hearing, 2000). In school-age children, acquired sensorineural hearing loss may occur as a result of disease or noise exposure. The effects of sensorineural hearing loss in children are well documented. Even mild, high-frequency, unilateral sensorineural hearing loss can have important developmental consequences (Bess, Dodd-Murphy, and Parker, 1998). More severe losses are likely to affect the development of speech, language, academic performance, and social-emotional development (Gallaudet University, 1998). In addition to sensorineural loss, hearing screening is needed to identify children with conductive hearing loss. In nearly all cases, conductive hearing loss in school-age children is due to otitis media. The incidence of otitis media with effusion (OME) is highest during the infant-toddler period and declines substantially by school age. Even so, otitis media is the most frequent primary diagnosis in children less than 15 years old (American Academy of Pediatrics, 1994). The hearing loss associated with OME, although mild and rarely permanent, can occur throughout childhood and may result in medical complications as well as potentially adverse effects on communication and development.

**Principles and Methods**

The purpose of hearing screening is to identify children most likely to have a hearing or middle ear disorder needing medical, audiologic, or other interventions. Thus, the goal of a hearing screening program is to identify asymptomatic individuals with an increased likelihood of hearing impairment, so that diagnostic follow-up is applied only to that subset of individuals. To justify the resources needed to provide a comprehensive screening program, several important assumptions must be met. The problem must be considered significant, both to individuals affected and to society; there must be appropriate screening tools with acceptable performance criteria; there must be effective treatment for those identified; and there must be sufficient financial resources for the program’s implementation and maintenance. Hearing screening in the school-age population satisfies each of these criteria; however, ongoing evaluation of new technologies and protocols is needed to determine optimal methodology and pass-fail criteria.

The Panel on Audiologic Assessment of the American Speech-Language-Hearing Association (ASHA, 1997a) recommends that school-age children be screened on initial school entry, annually from kindergarten through third grade, and in grades 7 and 11. The panel also recommends screening children at entry to special education, those who repeat a grade, and any who are newly admitted to the school system. More aggressive hearing screening is recommended for children with one or more of the high-risk indicators listed in Table 1.

<table>
<thead>
<tr>
<th>Table 1. School-Age Children in Need of Regular Hearing Screening and Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Parent/care provider, health care provider, teacher, or other school personnel have concerns regarding hearing, speech, language, or learning abilities</td>
</tr>
<tr>
<td>• Family history of late or delayed-onset hereditary hearing loss</td>
</tr>
<tr>
<td>• Recurrent or persistent otitis media with effusion for more than 3 months</td>
</tr>
<tr>
<td>• Craniofacial anomalies, including those with morphological abnormalities of the pinna and ear canal</td>
</tr>
<tr>
<td>• Stigmata or other findings associated with a syndrome known to include sensorineural or conductive hearing loss</td>
</tr>
<tr>
<td>• Head trauma with loss of consciousness</td>
</tr>
<tr>
<td>• Reported exposure to potentially damaging noise levels or ototoxic drugs</td>
</tr>
</tbody>
</table>

Guidelines and position statements of ASHA and the American Academy of Audiology (AAA) recommend visual inspection of the external ear to detect conspicuous signs of disease or malformation (AAA, 1997; ASHA, 1997b). The outer ear examination is followed by otoscopic inspection. Screening personnel must have the knowledge and skill required to conduct visual inspection and otoscopy, not for diagnostic purposes but to identify obvious signs of ear disease, impacted cerumen, or foreign objects that may compromise the validity of the screening or indicate the need for medical referral.

Current guidelines and position statements recommend that hearing screening of school-age children be conducted on an individual basis, using manually administered pure tones delivered via earphones or insert receivers at 20 dB hearing level (HL) for the frequencies 1000, 2000, and 4000 Hz (AAA, 1997; ASHA, 1997a). In order to pass, the child must respond to all three frequencies in each ear. Although failure to respond may simply be due to lack of cooperation or motivation, hearing loss should be suspected until appropriately ruled out. When referral is indicated, evaluation by an audiologist is needed to determine the nature and degree of hearing loss.

Routine pure-tone screening at 20 dB HL, however, is inadequate for the identification of OME (Melnick,
Eagles, and Levine, 1964; Roush and Tait, 1985). Consequently, many institutional screening programs include acoustic immittance (tympanometry and related measures) as part of the hearing screening program. Because OME has low incidence in the school-age population, ASHA guidelines recommend routine middle ear screening only to age 6 (ASHA, 1997b).

Evoked otoacoustic emissions (EOAEs) have become an essential tool in evaluating peripheral auditory function. EOAEs, which are present in nearly all ears with normal cochlear function, have at least two important advantages over behavioral pure-tone screening. They are objective and, for most school-age children, easy to measure. Because EOAEs are usually absent in ears with more than mild hearing loss, they are potentially well-suited for school-age screening (Driscoll, Kei, and McPherson, 1997; McPherson et al., 1998). But EOAEs are influenced by middle ear disease as well as by cochlear dysfunction. While this is often cited as a disadvantage of EOAEs, Nozza and colleagues have suggested that EOAEs, in conjunction with tympanometry, may be useful as part of a multistage screening program to detect hearing loss and otitis media (Nozza, Sabo, and Mandel, 1997; Nozza, 2001). Before such an approach can be endorsed for routine screening of school-age children, further clinical trials are needed to determine the sensitivity, specificity, and predictive value of EOAEs, Nozza and colleagues have suggested that EOAEs, in conjunction with tympanometry, may be useful as part of a multistage screening program to detect hearing loss and otitis media (Nozza, Sabo, and Mandel, 1997; Nozza, 2001).

Other Considerations

An ideal acoustical environment is rarely available in schools and other settings where screening is often conducted. Although few programs make on-site noise measurement as part of evaluating a test environment prior to screening, the time and expense involved in unnecessary follow-up and audiologic referral can far exceed the cost of providing a noise survey. Thus, a pre-screening noise survey that includes measurement of ambient noise levels at several third-octave bands (Table 2) should be conducted when there is uncertainty regarding adequacy of the screening environment (ANSI S3.1-1999).

Table 2. Maximum Permissible Ambient Noise Levels (in dB SPL) for One-Third Octave Bands, for Screening at 20 dB HL Using the Pure-Tone Frequencies 1000, 2000, and 4000 Hz (ANSI S3.1-1999)

<table>
<thead>
<tr>
<th>Stimulus/Transducer</th>
<th>Pure-Tone Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening at 20 dB HL with supra-aural earphones</td>
<td>41 49 52</td>
</tr>
<tr>
<td>Screening at 20 dB HL with insert earphones</td>
<td>62 64 65</td>
</tr>
</tbody>
</table>

Professionals and support personnel from many disciplines are now involved in hearing and middle ear screening. The screening procedures are not difficult with most school-age children; however, personnel must be competent and appropriately supervised. Furthermore, those responsible for the program must ensure that screening personnel are employed in a manner consistent with state licensure, professional scope of practice, and other regulatory requirements. For this reason it is recommended that institutional screening programs be conducted under the general supervision of an audiologist. School-age children who are difficult to test because of developmental disabilities or other factors should be screened by an audiologist.

Audiometers and tympanometric screening instruments must undergo full calibration by a qualified technician at least once each year. The American National Standards Institute (ANSI) has established specific requirements for the calibration of these instruments (ANSI S3.39-1987; S3.6-1996). In addition to formal calibration measurements, ANSI standards advise routine visual inspection and daily listening checks.

Because school-age children are at risk for permanent hearing loss from exposure to high-intensity noise, a comprehensive program of hearing screening for school-age children should include information on prevention of acoustic trauma through the use of appropriate hearing protection (Anderson, 1991). The New York League for the Hard of Hearing (2001) has developed useful materials for educating children about the dangers of noise-induced hearing loss and how to prevent it.

Parental permission must be obtained prior to conducting hearing and middle ear screening procedures unless consent has already been obtained as part of an institutional enrollment process or admissions procedure. Failure to obtain informed consent not only is unprofessional but could lead to negative public relations and possible legal action. Parents must be informed of screening procedures and their purpose. In addition to informed consent, strict confidentiality must be ensured. Discussion of screening outcomes or distribution of results should occur only with parents’ knowledge and consent, and mechanisms used to transmit screening results must be in secure data formats. Most screening programs will already have institutional guidelines. If guidelines do not exist, they must be implemented according to institutional protocols as well as state and federal laws. Screening personnel must be familiar with these policies and maintain compliance at all times (Roush, 2000).

Undetected hearing loss in a school-age child is a serious matter. Not only is appropriate intervention denied, hearing loss can be mistaken for a developmental disability or attention deficit. Even so, mass screening of school-age children is an expensive and time-consuming endeavor that requires systematic review and ongoing evaluation. This includes careful examination of screening outcomes, tracking of referrals, and communication with agencies and health care providers to whom referrals are made. In recent years, despite an enormous in-
crease in studies related to newborn hearing screening, there has been remarkably little research aimed at the school-age population. In particular, there is a need to further examine the role of EOAEs, alone and in combination with other tests, to determine the optimal screening battery for identification of hearing loss and middle ear disease in the school-age population.

—Jackson Roush

References


Further Readings


Hearing Protection Devices

Hearing protection devices (HPDs) are, as a practical matter, the first line of defense against hearing loss caused by excessive noise. Other ways to reduce exposure to loud sound (engineering control of noise sources, reduction of noise in the transmission path between a source and an individual) can indeed be more effective, but are often more costly or more difficult to manage.

HPDs can be classified by type (active versus passive), by form (e.g., earplugs and earmuffs of various types), and by effect. In all cases the goal is the same: to attenuate the magnitude of sound reaching the cochlea, thus limiting acoustical insult to the end-organ of hearing (see NOISE-INDUCED HEARING LOSS). Passive devices accomplish this through blockage of the airborne transmission path to the inner ear. Active devices seek to mechanically or electronically respond to noise to reduce signal amplitudes presented to the auditory system.

Earplugs fall into five categories (Berger, 2000): (1) closed-cell foam devices designed to be manually compressed, then inserted into the ear canal, where they expand to approximate their initial size (e.g., the Aearo
Company E-A-R plug), (2) preformed devices available in different diameters to accommodate different ear canals (e.g., the PlastiMed V-51R plug), (3) malleable devices intended to fit a range of ear canals, (4) semi-insert devices held in the ear canal by means of a plastic or metal band, and (5) devices made from ear mold impressions taken from individual ear canals. Most earplugs are created from plastics (polyvinyl chloride, polyurethane, silicone, or acrylic); malleable earplugs often consist of wax-impregnated cotton or fiberglass enclosed in a thin plastic container. Most earplugs are passive, that is, they are not intended to respond differently to differing noise exposures. Some active earplugs employ metal slugs or other material intended to move within the plug when stimulated by sudden acoustic overpressures, thus increasing attenuation in response to impulsive noise. For various reasons, it is difficult or impossible to objectively measure the attenuation of such devices or to estimate their real-world benefit. A recent addition to the array of earplugs are those offered by Etymotic Research for users with specific needs (e.g., musicians) who seek flat attenuation across specified frequency ranges.

Earmuffs are designed as integral components of safety helmets or as separate devices that surround the outer ear and are held in place by headbands that extend over or behind the head or beneath the chin. Some of these are combined with communication systems intended to increase the signal-to-noise ratio of messages electronically routed to earphones placed within headsets. At present, most earmuffs are passive devices, but several have been developed as active systems. Indeed, most active hearing protectors are based on earmuffs, owing to the space required for sound sensing and processing components. Active protectors employ one of two (or both) methods to attenuate sound. One method senses sound with a microphone, then processes sound delivered through an earphone by means of automatic gain control circuitry: when incident sound exceeds a certain level, further increases are electronically clipped or otherwise squelched. The other method samples incident sound, reverses the phase of the signal, and electronically adds the reversed signal within the muff enclosure to partially cancel the incident sound. Because of incident signal changes and processing speed requirements, devices employing additive cancellation techniques are more effective at relatively low frequencies (e.g., below 500 Hz; see Nixon, McKinley, and Steuver, 1992). Some active noise reduction methods appear similar to (or may benefit from) methods employed in hearing aids.

Beyond type and form, HPDs differ in weight, comfort, uniformity of fit to individuals, compatibility with other protective or prosthetic devices, and compatibility with individual user health status. Using eyeglasses with earmuffs, for example, can create acoustic leaks that reduce attenuation performance. Similarly, a subject with excessive cerumen or a middle ear effusion should not use earplugs. Use of hearing protectors in hot, humid environments can be uncomfortable and can cause skin irritation. If hearing protectors (perhaps combined with hearing loss) render speech communication difficult, or if they limit audibility of other signals deemed important, users may reject them. For obvious reasons, earmuffs should not be used in conjunction with hearing aids. These and other issues are discussed in detail by Berger (2000).

Real environments in which hearing protectors might provide benefit differ tremendously in noise amplitude, spectrum, and duration. Noise exposure is normally indexed by time-weighted average (TWA) sound pressure levels sampled using integrating meters or personal noise dosimeters. In the United States, such measurements are specified by Federal regulation (Occupational Safety and Health Administration [OSHA], 1983, CFR Part 1910.95) and the subject of technical standards (American National Standards Institute [ANSI], S12.19 1996). Among other details, exposure is to be indexed using a slow meter ballistic characteristic and an A-weighting network (a high-pass filter useful in predicting the effects of broadband noise on hearing). TWA levels are single-number values used to describe noise exposure and determine actions to protect workers from noise-induced hearing loss in the workplace (OSHA, 1983).

In 1979, the Environmental Protection Agency (EPA) issued a regulation intended to promote laboratory measurement of hearing protector attenuation for the purpose of combining such information with exposure data to estimate protective effect. The EPA regulation (CFR40 Part 211) built on previous technical standards (ANSI, 1974), and invoked a single-number index, called the Noise Reduction Rating (NRR), to be included in hearing protector product labels. Computation of NRRs from averaged behavioral real-ear attenuation-at-threshold (REAT) data assume temporally continuous band-limited noise stimuli with equal energy per octave (pink noise), and address intersubject variability by doubling the standard deviation of threshold shifts, then subtracting that value from mean threshold shift for each noise band. Adjusted attenuation values are summed logarithmically across stimulus noise bands to yield an NRR in decibels. Because the NRR method also assumes measurement of unprotected levels indexed with a C-weighting network (which has a flatter frequency response than the A-weighting network used to measure exposure), an additional 7 dB must be subtracted from the NRR to estimate A-weighted noise levels when a hearing protector is in place (see OSHA, 1983, Appendix B). Finally, because it was recognized that how hearing protectors are placed in a subject’s ears (plugs) or on a subject’s head (muffs) could affect outcomes, the EPA method specified experimenter fitting of HPDs during laboratory testing.

Because real-ear attenuation methods performed following the procedures stipulated by EPA designate experimenter fitting of HPDs, it is to be expected that resulting NRRs will be larger than what would be found with subject fitting of HPDs. Because all extant methods for measuring REATs use temporally continuous noise, results of such measurements cannot be generalized
to impulse noise (e.g., gunfire) or impact noise (e.g., forging).

Shortly after the inception of the current OSHA Hearing Conservation Rule (OSHA, 1983), the National Institute of Occupational Safety and Health (NIOSH) recommended that labeled NRRs be derated to estimate effectiveness in the field. Six schemes are noted in Appendix B of the Hearing Conservation Rule. These differ based on available measurement devices and data, but generally reduce the estimated benefit of hearing protectors. For example, if only A-weighted noise exposure data are available, 7 dB is subtracted from the NRR. For both A-weighted and C-weighted exposure data, the resulting corrected NRR is further reduced by 50%.

Subsequent research over two decades suggests that NRRs derated in this manner still overestimate the attenuation of hearing protectors in real-world situations. Various factors contribute to this inaccuracy, including overestimation associated with (1) experimenter fit, (2) highly trained test subjects (whose small standard deviations of REATs produce higher NRRs), and (3) differences in patterns of use of hearing protectors in laboratory and field settings (NIOSH, 1998).

Other pertinent generalizations include the following: (1) overall, earmuffs provide the most protection, foam and formable earplugs provide the next greatest protection, and all other insert types provide less, and (2) ideally, individuals should be fitted individually for hearing protectors (NIOSH, 1998). Generally, both earplugs and earmuffs provide greater attenuation at frequencies below 500 Hz than at lower frequencies (Berger, 2000).

Chapter 4 of the revised NIOSH criteria document (NIOSH, 1998) offers details about estimated real-world NRRs for 84% of wearers of hearing protectors, based on several independent studies. Labeled NRRs for single protectors range from 11 to 29 dB, while weighted mean NRR84 values range from 0.1 to 14.3 dB.

To address these problems, existing standards for measuring HPD attenuation were revised (Royster et al., 1996; Berger et al., 1998) to include subject-fit methods with audiometrically proficient listeners naive about HPDs. The resulting standard is ANSI S12.6 (1997). (A companion standard, ANSI S12.42 (1995), specifies a test fixture method and a microphone-in-real-ear method for measuring insertion loss useful for quality control and product development work with earmuffs.)

In 1995 the National Hearing Conservation Association proposed alternative labeling requirements in which only subject-fit real-ear attenuation data (ANSI S12.6-1997, Method B) are reported. The revised NRR(SF) information generally suggests less protection than NRRs based on experimenter fitting. Alternatively, the NHCA (1995) suggests labeling to include high, medium, and low NRRs based on statistical distributions of measured subject-fit REATs. This proposal has been endorsed by several other organizations. As of this writing, however, the EPA NRR labeling requirement remains based on the experimenter-fit method specified in ANSI S3.19 (1974).

If only experimenter-fit data are available, NIOSH (1998) currently recommends derating of NRRs based on type of hearing protector: 25% for earmuffs, 50% for formable earplugs, and 70% for all other earplugs. In the case of double protection (plugs and muffs), the OSHA Technical Manual (OSHA, 1999) recommends using the EPA NRR for the better protector, minus 7 dB, dividing the result by 2 (a 50% derating), then adding 5 dB to the field-adjusted NRR to account for the second protector.

Rather clearly, much work remains to be done to improve the prediction of real-world benefit of hearing protectors (Berger and Lindgren, 1992; Berger, 1999). One promising approach involves methods similar to in vivo real-ear gain measurements of hearing aids (now a common practice), together with modification of commonly used personal noise dosimeters. This approach requires the ability to simultaneously measure exposure level and the sound level generated within the ear canal of the wearer of a hearing protector. If both are measured with the same filtering schemes (preferably, the C-weighting network; ideally with both A and C networks), the signed difference between the two would index attenuation due to the hearing protector. If such measurements can be adapted to field use (e.g., with a two-channel noise dosimeter), it may be possible to add useful information to what otherwise can be determined about the performance of at least some HPDs. ANSI S12.42 (1995) addresses some of these issues for earmuffs and communication headsets, but only for laboratory measurements. Because the use of probe microphones with earplugs is likely to produce reactive measurement effects, this approach may not be suitable for insert devices.

It is generally recognized that effective use of hearing protectors in the workplace or elsewhere is influenced by factors that go beyond the physical performance of these devices. As summarized by NIOSH (1998), these factors include convenience and availability, comfort and ease of fit, compatibility with other safety equipment, and worker belief that the device can be worn effectively, will indeed prevent hearing loss, and will still permit hearing of important sounds.

—Michael R. Chial

References


Further Readings

Masking
A major goal of the basic audiologic evaluation is assessment of auditory function of each ear. There are situations during both pure-tone and speech audiometry, however, when the non-test ear can contribute to the observed response from the test ear. Whenever it is suspected that the non-test ear is responsive during evaluation of the test ear, a masking stimulus must be applied to the non-test (i.e., contralateral) ear in order to eliminate its participation.

Cross-hearing occurs when a stimulus presented to the test ear “crosses over” and is perceived in the non-test ear. It is the result of limited interaural attenuation during both air- and bone-conduction testing. Interaural attenuation refers to the reduction of energy between ears. Generally, it represents the amount of separation between ears during testing. Specifically, it is the decibel difference between the hearing level of the signal at the test ear and the hearing level reaching the non-test ear cochlea.

A major factor that affects interaural attenuation is the transducer type: air-conduction versus bone-conduction. Two types of earphones are commonly used during air-conduction audiometry. Supra-aural earphones use cushions that press against the pinna, while insert earphones are coupled to the ear by insertion into the ear canal.

Interaural attenuation for supra-aural earphones varies across frequency and subject, ranging from about 40 dB to 80 dB (e.g., Coles and Priede, 1970; Snyder, 1973; Killion, Wilber, and Gudmundsen, 1985; Sklare and Denenberg, 1987). The smallest reported value of interaural attenuation for speech is 48 dB (e.g., Snyder, 1973; Martin and Blythe, 1977). When making a decision about the need for contralateral masking during clinical practice, a single value defining the lower limit of interaural attenuation is most useful (Studebaker, 1967). The majority of audiologists use an interaural attenuation value of 40 dB for all air-conduction measurements,
both pure-tone and speech, when making a decision about the need for contralateral masking (Martin, Champlin, and Chambers, 1998).

Commonly used insert earphones are the ER-3A (Etymotic Research, 1991) and the E-A-RTONE 3A (E-A-R Auditory Systems, 1997). A major advantage of the 3A insert earphone is increased interaural attenuation for air-conducted sound, particularly in the lower frequencies. Consequently, the need for contralateral masking is significantly reduced during air-conduction audiometry. Based on currently available data, conservative estimates of interaural attenuation for 3A insert earphones with deeply inserted foam eartips are 75 dB at 1000 Hz and below and 50 dB at frequencies above 1000 Hz (Killion, Wilber, and Gudmundsen, 1985; Sklare and Denenberg, 1987). The smallest reported value of interaural attenuation for speech is 20 dB greater when using 3A insert earphones with deeply inserted foam eartips (Sklare and Denenberg, 1987) than when using a supra-aural arrangement (Snyder, 1973; Martin and Blythe, 1977). Consequently, a value of 60 dB represents a conservative estimate of interaural attenuation for speech when using 3A insert earphones.

Interaural attenuation is greatly reduced during bone-conduction audiometry. Regardless of the placement of a bone vibrator (i.e., mastoid versus forehead), it is generally agreed that interaural attenuation for bone-conducted sound is negligible and should be considered 0 dB (e.g., Hood, 1960; Sanders and Rintelmann, 1964; Studebaker, 1967; Dirks, 1994).

When to Mask

Contralateral masking is required during pure-tone air-conduction audiometry when the unmasked air-conduction threshold obtained in the test ear (ACT) minus the apparent bone-conduction threshold (i.e., the unmasked bone-conduction threshold) in the nontest ear (BCNT) equals or exceeds interaural attenuation (IA):

\[ \text{ACT} - \text{BCNT} \geq \text{IA} \]

Many audiologists will obtain air-conduction thresholds prior to measurement of bone-conduction thresholds. A preliminary decision about the need for contralateral masking can be made by comparing the air-conduction thresholds of the two ears. When the air-conduction threshold in the test ear minus the air-conduction threshold in the nontest ear (ACNT) equals or exceeds interaural attenuation, masking should be applied to the nontest ear:

\[ \text{ACNT} \geq \text{IA} \]

It is important to remember, however, that cross-hearing for air-conducted sound occurs primarily through the mechanism of bone conduction. Consequently, it will be necessary to reevaluate the need for contralateral masking during air-conduction testing following the measurement of unmasked bone-conduction thresholds.

The major factor to consider when making a decision about the need for contralateral masking during bone-conduction audiometry is whether the unmasked bone-conduction threshold (unmasked BC) suggests the presence of a significant conductive component in the test ear. Specifically, the use of contralateral masking is indicated whenever the results of unmasked bone-conduction audiometry suggest the presence of an air-bone gap in the test ear (AB GapT) of 15 dB or greater:

\[ \text{AB GapT} \geq 15 \text{ dB} \]

where

\[ \text{AB GapT} = \text{ACT} - \text{Unmasked BC} \]

Contralateral masking is indicated during speech audiometry whenever the presentation level of the speech signal in dB HL at the test ear (PLT) minus interaural attenuation equals or exceeds the best pure-tone bone-conduction threshold in the nontest ear (Best BCNT):

\[ \text{PLT} - \text{IA} \geq \text{Best BCNT} \]

Because speech is a broadband signal, it is necessary to consider bone-conduction hearing sensitivity at more than a single pure-tone frequency. The most conservative approach involves considering the best bone-conduction threshold in the 250- to 4000-Hz frequency range (Coles and Priede, 1975).

Clinical Masking Procedures

Although there are many different approaches to clinical masking, each addresses two basic questions. First, what is the minimum level of noise that is required to just mask the cross-hearing signal in the nontest ear? Stated differently, this is the minimum masking level that is needed to prevent undermasking (i.e., the test signal continues to be perceived in the nontest ear). Second, what is the maximum level of noise that can be presented to the nontest ear that will not shift or change the true threshold in the test ear? Stated differently, this is the maximum masking level that can be used without overmasking.

Masking refers to “the process by which the threshold of hearing for one sound is raised by the presence of another (masking) sound” (ANSI S3.6-1996, p. 5). The purpose of contralateral masking is to reduce the sensitivity of the nontest ear to the test stimulus. The masking noise typically used during pure-tone audiometry is narrow-band noise centered geometrically around the audiometric test frequency. Speech spectrum noise (i.e., weighted random noise for the masking of speech) is typically used during speech audiometry. Masking noise is calibrated in effective masking level (dB EM) (ANSI S3.6-1996). Effective masking level for pure tones refers to the dB HL to which detection threshold is shifted by a given level of noise. Effective masking level for speech refers to the dB HL to which the speech recognition threshold (SRT) is shifted by a given level of noise.
The introduction of contralateral masking can produce a small threshold shift in the test ear even when masking level is insufficient to produce overmasking. Wegel and Lane (1924) referred to this phenomenon as *central masking*. Central masking has been reported to affect thresholds during both pure-tone and speech audiometry (e.g., Lidén, Nilsson, and Anderson, 1959; Studebaker, 1962; Dirks and Malmquist, 1964; Martin, Bailey, and Pappas, 1965; Martin, 1966; Martin and DiGiovanni, 1979). Although the threshold shift generally is considered to be about 5 dB, variable results have been reported across subjects and studies.

The most popular method for obtaining masked pure-tone threshold was first described by Hood in 1957 (Hood, 1960; Hood’s original paper was reprinted in the United States in 1960). The Hood method is also referred to as the plateau, shadowing, or threshold shift procedure. The goal of the plateau procedure is to establish the hearing level at which the pure-tone threshold remains unchanged with increments in masking level. The masking “plateau” represents a range of masking levels (e.g., 15–20 dB) over which the pure-tone threshold remains unchanged. The recommended clinical procedure is summarized below:

1. Masking noise is introduced to the nontest ear at a minimum masking level. The pure-tone threshold is then reestablished.
2. The level of the tone or noise is increased subsequently by 5 dB. If there is a response to the tone in the presence of the noise, the level of the noise is increased by 5 dB. If there is no response to the tone in the presence of the noise, the level of the tone is increased in 5-dB steps until a response is obtained.
3. A plateau has been reached when the level of the noise can be increased over a range of 15–20 dB without shifting the threshold of the tone. This corresponds to a response to the tone at the same hearing level at three to four consecutive masking levels.
4. The masked pure-tone threshold corresponds to the hearing level of the tone at which a masking plateau has been established.

Formulas have been proposed for the calculation of minimum masking level during pure-tone threshold audiometry (e.g., Lidén, Nilsson, and Anderson, 1959; Studebaker, 1964). The simplified method described by Martin (1967, 1974) is recommended for clinical use. Specifically, minimum masking level ($\text{M}_{\text{Min}}$) (i.e., “initial masking” level) during air-conduction testing is equal to the air-conduction threshold of the nontest ear plus a safety factor of at least 10 dB:

$$\text{M}_{\text{Min}} = \text{AC}_{\text{NT}} + 10 \text{ dB}$$

Minimum masking level during bone-conduction audiometry is determined similarly. However, it is also necessary to account for the occlusion effect (OE):

$$\text{M}_{\text{Min}} = \text{AC}_{\text{NT}} + \text{OE} + 10 \text{ dB}$$

Whenever an earphone covers or occludes the nontest ear during masked bone-conduction audiometry, an occlusion effect may be created in the nontest ear. The nontest ear consequently can become more sensitive to bone-conducted sound for test frequencies below 2000 Hz, particularly when using supra-aural earphones (e.g., Elpern and Naunton, 1963; Goldstein and Hayes, 1965; Berger and Kerivan, 1983; Dean and Martin, 2000). Consequently, the minimum masking level must be increased by the amount of the occlusion effect. There is evidence suggesting that the occlusion effect is decreased significantly when using deeply inserted E-A-R foam plugs, the eartips used with 3A insert earphones (Berger and Kerivan, 1983; Dean and Martin, 2000). The clinician can use either individually determined (Martin, Butler, and Burns, 1974; Dean and Martin, 2000) or fixed occlusion effect values (i.e., based on average data reported in the literature) when calculating minimum masking level. When using supra-aural earphones, the following fixed occlusion effect values are recommended for clinical use: 30 dB at 250 Hz, 20 dB at 500 Hz, and 10 dB at 1000 Hz. When using 3A insert earphones with deeply inserted foam eartips, the following values are recommended: 10 dB at 250 and 500 Hz, and 0 dB at frequencies of 1000 Hz and higher. It should be noted that the occlusion effect is decreased or absent in ears with conductive hearing impairment (e.g., Martin, Butler, and Burns, 1974; Studebaker, 1979). If the nontest ear exhibits a potential air-bone gap of 20 dB or more, the occlusion effect should not be added to the minimum masking level.

The American Speech-Language-Hearing Association (ASHA, 1990) has published guidelines for audiometric symbols and procedures for graphic representation of frequency-specific audiometric findings. Different symbols are used to represent pure-tone thresholds obtained with and without contralateral masking. The reader is referred to ASHA’s 1990 guidelines and the article Pure-Tone Threshold Assessment.

The optimal masking level during speech audiometry is one that falls above the minimum and below the maximum masking levels (Lidén, Nilsson, and Anderson, 1959; Studebaker, 1979; Konkle and Berry, 1983). The goal is to select a masking level that falls at the middle of the masking plateau. This concept was originally discussed by Luscher and König in 1955 (cited in Studebaker, 1979).

Minimum masking level ($\text{M}_{\text{Min}}$), adapted from Lidén, Nilsson, and Anderson (1959), can be summarized using the following equation:

$$\text{M}_{\text{Min}} = \text{PL}_T - \text{IA} + \text{Max AB Gap}_{\text{NT}}$$

$\text{PL}_T$ represents the presentation level of the speech signal in dB HL at the test ear, $\text{IA}$ is the interaural attenuation value for speech, and $\text{Max AB Gap}_{\text{NT}}$ is the maximum air-bone gap in the nontest ear in the 250–4000-Hz frequency range. $\text{PL}_T - \text{IA}$, an estimate of the hearing level of the speech signal that has reached the nontest ear, represents the minimum masking level required. However, the presence of air-bone gaps in the nontest ear will reduce the effectiveness of the masker. Consequently, the
minimum masking level must be increased by the size of the air-bone gap. Following the recommendation of Coles and Prie(1975), the maximum air-bone gap in the non-test ear should be considered when determining minimum masking level.

Maximum masking level (M_{Max}), adapted from Lidén, Nilsson, and Anderson (1959), can be summarized using the following equation:

\[
M_{Max} = \text{Best BC}_T + IA - 5 \text{ dB}
\]

Best BC_T represents the best bone-conduction threshold in the test ear in the frequency range from 250 through 4000 Hz, and IA is the interaural attenuation value for speech. There is the assumption that the best bone-conduction threshold is most susceptible to the effects of overmasking. If Best BC_T + IA is just sufficient to produce overmasking, then a slightly lower masking level than the calculated value must be used clinically. Because masking level is typically adjusted using a 5-dB step size, a value of 5 dB is subsequently subtracted from the calculated value.

Although Studebaker (1962) originally described an equation for calculating midmasking level during pure-tone bone-conduction testing, the basic principles underlying the midplateau method also can be applied effectively during speech audiometry. A direct approach to calculating the midmasking level (M_{Mid}) involves determining the arithmetic mean of the minimum and maximum masking levels:

\[
M_{Mid} = \frac{(M_{Min} + M_{Max})}{2}
\]

Yacullo (1999) has described a simplified approach to selecting an appropriate level of contralateral masking during speech audiometry. Although this approach proves most effective during assessment of suprathreshold speech recognition, it can also be applied during threshold measurement. Stated simply, effective masking level is equal to the presentation level of the signal in dB HL at the test ear minus 20 dB:

\[
\text{dB EM} = \text{PL}_T - 20 \text{ dB}
\]

The procedure proves very effective given the following two prerequisite conditions: (1) there are no significant air-bone gaps (i.e., ≥15 dB) in either ear, and (2) speech is presented at a moderate sensation level (i.e., 30–40 dB) relative to the measured or estimated SRT. Given these two prerequisites, the selected masking level will occur approximately at midplateau.

The plateau masking procedure, described earlier as a popular method for obtaining masked pure-tone threshold, also can be applied effectively during measurement of both speech detection and speech recognition thresholds (i.e., SDT and SRT). A major advantage of the plateau procedure is that information about bone-conduction sensitivity in each ear is not required when selecting appropriate masking levels. The reader is referred to Yacullo (1996) for further discussion.

—William S. Yacullo

References


**Further Readings**


**Middle Ear Assessment in the Child**

Current clinical methods of assessment of the middle ear in children include otoscopic examination, acoustic immittance measures, and reflectometry. When assessing middle ear function, a good otoscopic evaluation is the first step. Examination of the ear canal for any obstructions that would preclude placement of a probe such as is used for acoustic immittance measures is essential. Often, cerumen in the ear canal becomes impacted, even in children, and thereby confounds tympanometric measures. Even when not impacted, cerumen can clog the immittance probe and cause invalid measurements.

Otoscopy can also be useful in identifying middle ear disorders such as middle ear effusion. Examination of the tympanic membrane can provide evidence of fluid by its opacity or color. An opaque or yellow membrane, for example, might indicate middle ear effusion. Clearer evidence comes when a fluid meniscus or bubbles can be seen through a transparent tympanic membrane. Someone skilled with pneumatic otoscopy can examine the membrane’s mobility by applying slight changes in air pressure in the ear canal. However, the ability to use a pneumatic otoscope for the diagnosis of middle ear effusion is highly variable. Of course, in the clinical examination associated with the diagnosis of middle ear effusion, visual examination of the ear canal and tympanic membrane is important so that other conditions (e.g., cholesteatoma, retraction pocket) may be identified.

**Acoustic Immittance: Tympanometry**

Tympanometry, unlike pneumatic otoscopy, is a derived physiological measure that requires instrumentation that meets a standard of the American National Standards Institute (ANSI S3.39, 1987) and is a test that is easily administered. The principles of acoustic immittance measurement are covered elsewhere.

Acoustic immittance is a term used to describe the ability of energy to flow through the middle ear. The word immittance is a combination of the words impedance (opposition to the flow of energy) and admittance (ease with which energy flows). Today, most immittance instruments measure acoustic admittance, so the description that follows will consider only acoustic admittance measures. *Tympanometry* is a dynamic measurement of acoustic admittance as air pressure in the ear canal is varied. The tympanogram is the graph of acoustic admittance in mmhos (or in equivalent volume units such as milliliters) versus air pressure in decapascals (daPa). To estimate admittance, a tone is delivered to the ear via a probe hermetically sealed in the ear...
canal. The probe tone is the force that activates the middle ear system, while a microphone that is connected to a separate opening in the probe records sound pressure in the ear canal. A third opening in the probe connects to a pneumatic system for varying air pressure in the ear canal.

The first measurement made in tympanometry is an estimate of the volume of the space between the probe tip and the tympanic membrane and is called ear canal volume or equivalent volume. This is done with a low-frequency (226-Hz) probe tone when the pressure in the ear canal is set to a very high value (either positive or negative). The extreme pressure stiffens the tympanic membrane such that admittance of the middle ear is diminished to (theoretically) zero. Thus, the acoustic admittance detected by the instrument is a measure of the admittance of the volume of air enclosed in the ear canal and, with the 226-Hz probe tone only, can be expressed as an equivalent volume of air in the ear canal. This ear canal volume measure is useful for determining the validity of the probe fit in the ear canal and, when larger than normal, suggests an opening in the tympanic membrane. This is useful when trying to determine if there is a perforation in the tympanic membrane or if a pressure equalization tube is functioning.

The normal tympanogram has the shape of an inverted V, with the peak amplitude referred to as peak admittance or static admittance. When the admittance attributed to the ear canal volume is automatically subtracted from the dynamic admittance measure, peak admittance is referred to as peak compensated acoustic admittance. The magnitude of the peak in mhmhos of admittance carries diagnostic information. Because there is developmental change in acoustic admittance variables, peak admittance in a given case must be compared with age-appropriate normative values. Low peak admittance when tympanometric peak pressure (the air pressure in the ear canal at which the peak occurs) is within normal range suggests that the middle ear space is well aerated but that there is a condition that is reducing the ability of sound to flow through. A condition that increases stiffness in the middle ear system, such as stapes fixation, might result in such a pattern.

If peak admittance is greater than the normal range, with tympanometric peak pressure within the normal range, there is a condition of the middle ear that is increasing admittance, such as an ossicular disarticulation. Because middle ear function is being inferred from admittance changes based on sound reflected from the tympanic membrane, abnormalities of that membrane can influence the measurement. For example, scarring on the membrane can also cause an abnormally high admittance tympanogram in the presence of an otherwise normal middle ear system.

In an ear with eustachian tube dysfunction, in which middle ear pressure varies and is not equalized easily, tympanometry can give valuable information. The tympanometric peak pressure is a reasonably good estimate of middle ear pressure, so a negative value is often diagnostic of eustachian tube dysfunction.

**Pattern Classification of Tympanograms**

Historically, the most common way to interpret tympanometric data has been to classify tympanograms according to their pattern. The most notable and widely used pattern classification scheme was proposed by Jerger (1970). This pattern classification scheme is easy to use, with patterns that are easy to identify and few in number. The pattern classifications were based originally on the height of the tympanogram (in arbitrary “compliance” units) and the tympanometric peak pressure. The patterns are identified by the letters A, B, and C, with subclassifications used to better define them. A normal tympanogram is A, with A1 and A2, representing abnormally low peak compliance and abnormally high peak compliance, respectively, in the presence of normal tympanometric peak pressure. Type B is a tympanogram with no peak, and type C is a tympanogram with a high negative tympanometric peak pressure.

In children, the most common reason for assessment of middle ear function is to identify middle ear effusion. As a result, much of the published research on tympanometry in children relates to identification of middle ear effusion. In most studies of the ability of tympanometry to identify ears with middle ear effusion, the type B tympanogram (no peak) has high predictive ability; that is, a tympanogram that has no peak will very likely be associated with an ear with effusion. However, the opposite is not necessarily true: ears with middle ear effusion most often produce tympanograms that cannot be classified as type B. Rarely is an ear with middle ear effusion associated with a type A (normal peak, normal tympanometric peak pressure), so many ears with middle ear effusion fall in the C category (normal peak admittance, abnormally negative tympanometric peak pressure). However, many ears with no middle ear effusion also fall into the C category. Subdivisions of the C category have emerged to try to improve the diagnostic ability of tympanometry, but even with more specific information regarding tympanometric peak pressure, the ability of the C category to help with identification of middle ear effusion varies widely across studies (Orchik, Dunn, and McNutt, 1978).

Because of the ambiguity in the C category, even when subdivided, attempts to further improve identification of middle ear effusion have included addition of the acoustic reflex test, in which case an equivocal tympanometric pattern accompanied by an absent acoustic reflex is considered evidence of middle ear effusion. Also, schemes with very detailed categories have been developed (e.g., Cantekin et al., 1980).

There are several drawbacks to the A, B, C pattern classification scheme. One is that the patterns were determined based on clinical observations and not on statistical analysis of performance in a controlled clinical trial with a reference standard such as myringotomy for the diagnosis of middle ear effusion. Many studies that have been done to determine the performance characteristics of the pattern classification of tympanograms used predetermined categories rather than collecting
Quantitative Analysis of Tympanograms

Instruments developed since 1987, when the ANSI standard became effective, provide immittance information using absolute physical quantities rather than arbitrary compliance values such as were used when the pattern classification scheme was developed. There are several advantages to using a quantitative analysis of tympanometric data. First, such measures are standardized for all instruments and, therefore, across clinics and laboratories. Second, tympanometric shape can be quantified. Third, unlike the pattern categories, the data from current instruments that measure actual physical quantities of admittance are on continua that permit statistical analyses. Measures that may be used in a quantitative analyses include peak compensated acoustic admittance, tympanometric peak pressure, absolute gradient, relative gradient, and tympanometric width.

The absolute and relative gradients and tympanometric width are measures used to quantify the shape of the tympanogram. To compute absolute gradient, a horizontal line is drawn between the sides of the tympanogram at the point where the tympanogram is 100 daPa wide. This serves as a temporary baseline from which the distance to the peak of the tympanogram is measured in mmoH. This value is the absolute gradient and increases in value as tympanogram shape becomes more sharply peaked. Relative gradient is determined by dividing the absolute gradient by the peak admittance. This results in a ratio that can range from 0 to 1. Again, the greater the relative gradient, the more sharply peaked is the tympanogram. Few instruments report absolute gradient, but some will report relative gradient to characterize the rate of change of the tympanogram in the region of the peak (i.e., tympanogram shape). Absolute and relative gradients have been examined in the past for their contribution to the diagnosis of middle ear effusion (Brooks, 1968). The lower the gradient, the more rounded or flat the tympanogram and the greater the likelihood of middle ear effusion.

Tymanometric width is the width of the tympanogram in daPa at half the peak admittance. The more rounded or flat the tympanogram becomes, the greater will be the tympanometric width. Tymanometric width has good diagnostic value for identification of middle ear effusion in children (Margolis and Heller, 1987; Nozza et al., 1994). Using quantitative analysis, Nozza et al. (1994) reported that using tympanometric width greater than 275 daPa as a criterion for identification of middle ear effusion had a sensitivity of 81% and a specificity of 82% in a group of children undergoing myringotomy and tube surgery. This was the best performance of any single tympanometric variable. For peak admittance, a cutoff of 0.3 mmoH separated the ears, with a sensitivity and specificity nearly as good as the best cutoff for tympanometric width. Interestingly, tympanometric peak pressure alone was the worst at separating ears with and without middle ear effusion and contributed nothing when used in combination with other variables. This suggests that the weight that tympanometric peak pressure carries in the pattern classification scheme is probably not warranted when it comes to identification of ears with middle ear effusion. The acoustic reflex alone was examined as well, and overall performance was only fair because specificity was poor. Too often, no acoustic reflex can be measured in children, even in the absence of middle ear effusion. Also, because the reflex relies on an auditory system and elements of the central nervous system that are sufficiently intact, an immittance criterion for identification of middle ear effusion that includes the acoustic reflex will not be applicable to children with high degrees of hearing impairment or certain neurological problems.

Because the data from studies such as that of Nozza et al. (1994) come from children from a special population, those undergoing myringotomy and tube surgery, they may not be representative of the general pediatric population, in whom middle ear assessment is so important. Recommendations for criteria for identification of middle ear effusion in the general population can be found in guidelines for screening that have been published by the American Academy of Audiology (1997) and the American Speech-Language-Hearing Association (ASHA, 1997). The American Academy of Audiology suggests using tympanometric variables of peak admittance or tympanometric width. In both guidelines, the notion is that a positive result on tympanometric screening should not result in immediate referral but should indicate the need for rescreening in 6 to 8 weeks. Because middle ear effusion is a transient, often self-limiting disorder, referrals based on a single test result in a high over-referral rate. ASHA recommends that an ear pass the tympanometric screen when peak admittance is ≥0.3 mmoH or tympanometric width is <200 daPa when screening children from the general population for middle ear effusion. The American Academy of Audiology uses a similar protocol, with ≥0.2 mmoH a passing indication for peak admittance and <250 daPa a passing indication for tympanometric width.
High-Frequency Tympanometry in Infants

With universal newborn hearing screening increasing the number of young infants referred for rescreening and audiological assessments, it is important to consider the special circumstances related to middle ear assessment in that population. It has long been known that tympanometry using low-frequency probe tones is unreliable in young infants. Infants with middle ear effusion can produce tympanograms that appear normal, presumably because of distensibility of the walls of the ear canal. Recent work using multifrequency tympanometry has demonstrated that young infants (<4 months) with middle ear effusion might produce a normal tympanogram with a 226-Hz probe tone but an abnormal tympanogram when a high-frequency probe, 1000 Hz in particular, is used (McKinley, Grose, and Roush, 1997; Rhodes et al., 1999; Purdy and Williams, 2000). Use of a 1000-Hz probe tone is recommended now for identification of middle ear effusion in young infants.

Reflectometry

An alternative method for identification of middle ear effusion uses a measurement called acoustic reflectometry. The acoustic reflectometer generates a broad band sound in the ear canal and measures the sound energy reflected back from the tympanic membrane. The instrument is hand-held and has a speculum-like tip that is put into the entrance of the ear canal. No hermetic seal is required, thus making the test desirable for use with children. The relationship between the known output of the device and the resultant sound in the ear canal provides diagnostic information. An early version of the instrument used only the sound pressure level in the ear canal in the diagnostic decision. It was later determined that, by plotting out the reactivity date on frequency by amplitude axes, better diagnostic information regarding middle ear effusion could be derived. The current version of the instrument analyzes automatically the frequency-amplitude relationship in the reflected sound and displays a number between 1 and 5 to indicate the likelihood of an effusion. This instrument is getting some favorable use in primary care settings for the identification of both acute otitis media and asymptomatic otitis media with effusion.

See also HEARING LOSS SCREENING: THE SCHOOL-AGE CHILD; OTITIS MEDIA: EFFECTS ON CHILDREN’S LANGUAGE; PEDIATRIC AUDIOLOGY: THE TEST BATTERY APPROACH; PHYSIOLOGICAL BASES OF HEARING; TYPMANOMETRY.

—Robert J. Nozza

References


Further Readings


Noise-Induced Hearing Loss

Hearing loss affects about 28 million Americans. The two most common causes of sensorineural hearing loss are aging and exposure to noise. Noise exposure can injure the ear and produce loss of hearing. The injury and hearing loss can be temporary—that is, fully recoverable after the noise exposure is terminated—or permanent. When hearing loss is measured at a postexposure time of 2 weeks, it is considered to be permanent, inasmuch as very little additional recovery occurs at postexposure times in excess of 2 weeks (Miller, Watson, and Covell, 1963; Mills, 1973). The presence and severity of a temporary or permanent hearing loss depend on several factors and the susceptibility of the individual. Acoustically, the level (intensity), spectrum (frequency), and temporal properties (duration, intermittency, number) of the exposures are the most pertinent properties. Nonacoustic factors such as interactions with various medicines or drugs, eye color, smoking, sex, and other personal characteristics of an individual (except aging and hearing loss) are second-order effects, or are not consistently observed (Ward, 1995). Interactions with other factors such as chemicals, solvents, and toxic substances such as carbon monoxide can be significant and are an active area of investigation (Morata and Dunn, 1995).

The database available on noise-induced hearing loss caused by exposure to steady-state noise is substantial. Some of the data are from laboratory studies of temporary effects in human subjects (Davis et al., 1950; Mills et al., 1970; Melnick and Maves, 1974) and both temporary and permanent effects in laboratory animals (Miller, Watson, and Covell, 1963; Mills, 1973). Other data are from studies of permanent hearing loss in humans in occupational settings (Taylor et al., 1965; Burns and Robinson, 1970; Johnson, 1991). These and other field studies of noise-induced permanent hearing loss (see Johnson, 1991) are the scientific bases of international (International Organization for Standardization [ISO], 1990) and American standards (American National Standards Institute [ANSI], 1996), which present methods to estimate noise-induced permanent threshold shifts as a function of A-weighted sound level and years of exposure time. The development and acceptance of these standards represents many years of work and intense debate. Regulations for industry are given by the U.S. Department of Labor (Occupational Safety and Health Administration [OSHA], 1983).

Data from humans and animals exposed to a wide variety of steady-state noises suggest that the range of human audibility can be categorized with respect to risk of acoustic injury of the ear and noise-induced hearing loss. This categorization is shown in Figure 1, where the range of human audibility is bounded by the threshold of...
audibility at one extreme and the threshold of pain (Yost and Nielsen, 1985) at the other. Of course, sounds below the threshold of audibility are inaudible and present no risk of noise-induced hearing loss. Sounds in excess of the threshold of pain present a risk of acoustic injury of the ear and noise-induced hearing loss even with one, short exposure (see Mills et al., 1993). Between the extremes of pain and audibility are acoustic injury thresholds (open triangles in Fig. 1). These thresholds define the highest levels of noise that will not produce a noise-induced threshold shift regardless of the duration of exposure, the number of exposures, or the temporal properties of the exposure. The data points in Figure 1 are from temporary threshold shift experiments for octave bands of noise from 63 Hz to 4 kHz in octave steps. Data at lower and higher frequencies are extrapolations. Thus, between the extremes of the threshold of pain and audibility are two categories: no risk and risk. The no-risk category can also be described as “effective quiet” (Ward, Cushing, and Burns, 1976). The region bounded by safe levels on the low side and threshold of pain on the high side is the area where the risk of hearing loss and acoustic injury of the ear depends on the parameters of the noise exposure as well as on the susceptibility of the individual. In qualitative terms, risk increases with noise level, duration, number of exposures, and susceptibility of the individual. Although individual differences can be substantial, no method has been developed that allows the a priori identification of those individuals who are most susceptible to noise-induced hearing loss. Quantitative relations between noise-induced hearing loss and exposure parameters are given in ANSI S3.44-1996.

Whereas the database is massive for noise-induced hearing losses produced by exposure to continuous noise (see Johnson, 1991), it is unimpressive for intermittent (quiet periods of a few seconds to a few hours) and time-varying (level fluctuations greater than 10 dB) noises. On a qualitative basis, there is agreement that intermittent and fluctuating noises are less hazardous than continuous noises, presumably because the “quiet” periods allow time for the ear to recover. Because of regulatory efforts in noise control and a perceived need for simplicity, several single-number correction factors have evolved. That is, as an exposure is increased from 4 hours to 8 hours, what change in noise level is needed to maintain an equal risk of hearing loss? The equal-energy rule specifies a 3-dB reduction in noise level for a doubling of exposure duration. This rule is incorporated into the ISO 1999 standard. Other standards and regulations use 4-dB, 5-dB, and 6-dB rules, as well as other more complicated schemes (see Ward, 1991). It is likely that each of these single-number rules may apply only to a restricted set of exposure conditions. With the continued absence of needed data, the effects of intermittence on noise-induced hearing loss may always be a contentious issue.

Although the biological bases of noise-induced hearing loss have been studied extensively, with the greatest emphasis on the cochlea, both the external and middle ear play a prominent role. The external auditory meatus (essentially a tube open at one end with a length of about 25 mm) has a resonant frequency of about 3 kHz and a gain of about 20 dB. Thus, the typical industrial or environmental noise, which may have a flat or slightly downward-sloping spectrum when measured in the field, will have a peak at 3 kHz because of the external ear canal. Thus, the “4-kHz notch” in the audiogram that is characteristic of noise-induced hearing loss (and head injuries) may reflect the acoustic properties of the external ear. In the middle ear, the acoustic reflex, the consensual contraction of the stapedial and tensor tympani muscles, may reduce the level of intense sounds. In addition, the efferent innervation of outer hair cells may have a protective role (Maison and Liberman, 2000).

Although there has been a substantial effort, the anatomical, chemical, and biological bases of noise-induced temporary threshold shift are unknown. The pathological anatomy associated with noise-induced permanent hearing loss involves the organ of Corti, especially the hair cells. Loss of outer hair cells is the most prominent anatomical feature of permanent noise-induced loss, and is almost always greater than the loss of inner hair cells. This greater loss of outer than inner hair cells may occur for several reasons, including the direct shearing forces on the outer hair cell stereocilia, which are embedded in the tectorial membrane. The correlations between loss of hair cells, both inner and outer, and permanent threshold shift are very high, ranging from 0.6 to 0.8, depending on frequency (Hamernik et al., 1989). Even with such high correlations there remains considerable variance between hair cell loss and permanent threshold shift. This variance can be reduced by consideration of the status of the stereocilia (Liberman and Dodds, 1984). With degeneration of inner hair cells following severe exposures, there can be retrograde degeneration of auditory nerve fibers, as indicated by losses of spiral ganglion cells. Neuronal degeneration is not restricted to the auditory nerve but progresses throughout the ascending auditory system (Morest, 1982). Regeneration of hair cells has been observed after intense exposures to noise. This dramatic effect has been reported only for the cochlea of various species of birds. Regenerating sensory cells have not been observed in the cochlea of mammals.

Reactive oxygen species and oxidative stress have been implicated in the production of noise-induced hearing loss and in age-related hearing loss as well (Ohlemiller et al., 2000). It is believed that acute impairment of antioxidant defenses promotes cochlear injury, and conversely, strengthening antioxidant defenses should provide protection, including possibly rescuing cells that are in the early stages of injury. Efforts at prevention and rescue are in the early stage of development, with some promising initial results (Hu et al., 1997). Additional protective functions can be obtained by conditioning exposures or exposures that protect the ear from subsequent noise (Canlon, Borg, and Flock, 1988). A related phenomenon is improvements (reductions) in threshold shifts observed during the course of an extended sequence of intermittent exposure to noise.
(Miller, Watson, and Covell, 1963; Clark, Bohne, and Boettcher, 1987). Clearly, noise-induced hearing loss is not related simply to the sound level of the exposure.

Acoustic trauma refers to injury of the ear and permanent hearing loss caused by exposure to an intense, short-duration sound. In contrast to the gradual loss of outer hair cells and stereocilia typically seen from steady-state or intermittent exposures with sound levels less than 100–110 dB SPL, the injury to the organ of Corti is more extensive, involving the tearing of membranes, rupturing of cells, and mixing of cochlear fluids. At extremely high sound levels, the tympanic membrane and middle ear can be injured, with a resultant conductive/mixed hearing loss. The most common form of acoustic trauma is hearing loss associated with the impulsive noises produced by small-arms gunfire (Clark, 1991). Hearing loss from impulses is related to the peak SPL of the impulse, duration, number of impulses, and other variables (Henderson and Hamernik, 1986; Hamernik and Hsueh, 1993; Hamernik et al., 1993).

A longstanding issue is the interaction between noise-induced hearing loss incurred throughout a person’s working lifetime and the hearing loss associated with aging. This issue is particularly important for medical reasons, for litigation involving worker’s compensation for occupational hearing loss, and in establishing noise standards and damage-risk criteria. Both the current ISO (1990) and ANSI (1996) standards assume that noise-induced permanent threshold shifts add (in decibels) to age-related threshold shifts, i.e., a 20-dB loss from noise and a 20-dB loss from aging results in a loss of 40 dB. Some data from noise-exposed and aged animals do not support additivity in decibels but additivity in intensity: i.e., 20 dB and 20 dB produces a loss of 23 dB (Mills et al., 1997). The issues of additivity and medical-legal aspects of noise-induced hearing loss are discussed by Dobie (2001).

—John H. Mills

References


Otoacoustic Emissions

The 25th anniversary of the discovery of otoacoustic emissions (OAEs), the acoustic energy produced by the cochlea, was celebrated in 2003. Following Kemp’s (1978, 1979a, 1979b) breakthrough descriptions of the four types of OAEs in humans—the click- or transient-evoked OAE (TEOAE), the distortion product OAE (DPOAE), the stimulus frequency OAE (SFOAE), and the spontaneous OAE (SOAE)—interest turned to basic research issues. Existing models of cochlear function were modified to reflect the existence of active processing as implied by the mere reality of OAEs. Also, efforts were made to relate OAEs to parallel neural and psychoacoustical phenomena, and to describe emitted responses in species used as research models, including monkeys, gerbils, guinea pigs, and chinchillas (Zurek, 1985).

During the early years of OAE study, another great advance in the hearing sciences occurred when Brownell et al. (1985) discovered electromotility in isolated outer hair cells. The current consensus is that outer hair cell motility is due to the receptor-potential initiated movements of atomic-sized “motor” molecules called prestin (Zheng et al., 2002) that are embedded in the lateral membrane of the outer hair cell. Our present understanding is that OAEs are generated as a by-product of these electromotile vibrations of outer hair cells (Brownell, 1990).

As the initial basic studies on OAEs were ongoing, the significant benefits of OAEs as a clinical test were being recognized. Thus, early on, four major applications of OAE testing in clinical settings became apparent: the differential diagnosis of hearing loss, hearing screening in difficult to test patients, serial monitoring of progressive hearing impairment conditions, and determining the legitimacy of medicolegal claims involving compensatory payments for hearing loss. The rationale for using OAEs in each of these major applications is based on a significant beneficial feature of the measure, including its specificity for testing the functional status of outer hair cells, the most fragile sensory receptors for hearing. This attribute in particular makes OAEs an ideal measure for determining the sensory component of a sensorineural hearing loss. In addition, mainly because...
the OAE is an objective response that is noninvasively measured from the outer ear canal and thus can be rapidly obtained, it is an ideal screening test for identifying hearing impairment in newborns. Finally, because OAEs are stable and reliably measured over long time intervals, they are excellent for monitoring pathological changes in cochlear function, particularly in individuals regularly exposed to ototoxic drugs or excessive sounds.

One relatively new application of OAEs over the past decade has been the use of emissions to measure the intactness of the entire ascending and descending auditory pathway (Collet et al., 1990). This capability is based on the knowledge that the suppressive effects of cochlear efferents mainly affect outer hair cell activity, since these sensory cells are the primary targets of the descending auditory system. Indeed, recent research indicates that the susceptibility of the ear to the harmful effects of, for example, intense noise is likely determined by the amount of indigenous efferent activity (Luebke, Foster, and Stagner, 2002). That is, the more robust the efferent activity, the more resistant the ear is to the damaging effects of loud sounds, and vice versa.

Of the two general classes of OAEs, SOAEs have not been as clinically useful as the evoked OAEs, for several reasons. Their prevalence in only about 50% of normal-hearing individuals and the individually based uniqueness of their frequencies and levels make it difficult to develop SOAEs into a standardized test. However, SOAEs have been linked to tinnitus in a subset of tinnitus patients with near-normal hearing (Penner, 1992). In patients with SOAE-induced tinnitus, suppressing the associated SOAEs eliminates the annoying tinnitus. Interestingly, and contrary to expectations, since excessive aspirin produces tinnitus in normal-hearing individuals, high-dose aspirin suffices as a palliative in persons with SOAE-induced tinnitus (Penner and Coles, 1992).

Concerning the three subclasses of evoked OAEs, only TEOAEs and DPOAEs have proved to be clinically useful. SFOAEs can be reliably measured only using expensive phase-tracking devices, since the emission must be extracted from the ear canal sound at a time that the eliciting stimulus is present at the identical frequency. Fortunately, the SFOAE is essentially the long-lasting version of the TEOAE, which is more straightforward to measure and interpret.

Within a decade after the discovery of TEOAEs, commercial equipment based on procedures used for evoking auditory brainstem responses was available (Bray and Kemp, 1987). Figure 1 shows the results of pure-tone audiometry (A) and tests of click-evoked counterpar distribution of noise-floor values for the same subjects. Note that the patient was exceptionally quiet, as his noise-floor curves (bold line = left ear; stippled line = right ear) tracked the lower distribution trajectory of the control population. In this example, the ototoxic drug caused outer hair cell dysfunction for frequencies above about 2.5 kHz for the left ear and 4 kHz for the right ear.
TEOAEs (B) and DPOAEs (C) in a 37-year-old patient receiving the ototoxic antitumor drug cisplatin. In this case, following a single infusion (Fig. 1A), a moderate high-frequency hearing loss was evident bilaterally for frequencies over 4 kHz. The associated TEOAE spectrum (Fig. 1B) illustrates the commonly measured properties for this emitted response, including its level, frequency content and extent, and reliability, according to automatically computed reproducibility factors for five representative frequencies at 1, 2, 3, 4, and 5 kHz.

Because the TEOAE is measured after the transient stimulus occurs, each ear produces a response that exhibits a unique spectral pattern. This idiosyncratic property makes it difficult to develop a set of metrics that describe the average TEOAE for normal-hearing individuals. Owing to this difficulty in determining “normal” TEOAEs in terms of frequencies and level values, they are most often described as being either present or absent. Thus, one of the most popular uses of TEOAEs clinically is as a test for screening auditory function in newborns (Norton et al., 2000).

In the example of Figure 1B, representing the left ear of the patient, even in the presence of a drug-induced high-frequency hearing loss, the TEOAE pattern appears fairly normal in that the click-elicited emission typically falls off for frequencies greater than 2 kHz, and is seldom present at frequencies above 4 kHz in adult ears. For newborns and older infants, the TEOAE is much more robust by about 10 dB and typically can be measured out to about 6 kHz, indicating that smaller ear canals influence the acoustic characteristics of standard click stimuli much differently than do adult ears.

Distortion product OAEs are elicited by presenting two long-lasting pure tonebursts at f1 (lower frequency) and f2 (higher frequency) simultaneously to the ear. The frequencies and levels of the tonebursts or primary tones are important in that the largest DPOAEs are elicited by f1 and f2 primaries that are within one-half octave of each other (i.e., f2/f1 = 1.22) with levels, L1 and L2, that are offset. For example, typical clinical protocols measure the 2f1–f2 DPOAE, which is the largest DPOAE in human ears, in response to primary-tone levels of L1 = 65 and L2 = 55 dB SPL (Gorga, Neely, and Dorn, 1999).

Figure 1C shows a DP-gram, i.e., DPOAE level as a function of test frequency, from about 800 Hz to 8 kHz, in response to equilevel primary tones (L1 = L2 = 65 dB SPL). In this example, test frequency is represented by the geometric or logarithmic mean of f1 and f2, although it could also be represented by the f2 frequency. That is, based on a combination of theoretical considerations, experimental studies, and observations of the generation of DPOAEs in pathological ears, it is clear that these emissions are produced in the region of the primary tones. Based on further experimental work, it is likely that the DPOAE source is level-dependent, with the primary generation site in response to higher level primaries of equal level (L1 = L2) occurring around the geometric mean frequency. In contrast, for lower level primaries, which are often offset in level, the primary generation site is closer to f2.

As illustrated in Figure 1C, the patient’s emissions were relatively normal, as compared to the ±1 SD distribution of DPOAE levels for normal-hearing adults, until about 3 kHz for the left ear (solid circles) and 4 kHz for the right ear (open circles). In this case, because DPOAEs are typically tested out to 8 kHz, they detected the developing high-frequency hearing loss associated with the ototoxic antitumor therapy.

It is clear that applications of OAEs in the hearing sciences and clinical audiology are varied. Without a doubt, OAEs are useful experimentally for evaluating and monitoring the status of cochlear function in animal models, and clinically in distinguishing cochlear from retrocochlear disorders. Moreover, their practical features make them helpful in the hearing screening of newborns. Additionally, they have proved useful in monitoring the effects of agents such as ototoxins and loud sounds on cochlear function. In fact, there is accumulating evidence that it is possible to detect such adverse effects of drugs or noise on outer hair cell function using OAEs before a related hearing loss can be detected by pure-tone audiometry. In addition, OAEs provide a noninvasive means for assessing the integrity of the cochlear efferent pathway. In general, OAEs supply unique information about cochlear function in the presence of hearing problems, and this capability makes them ideal response measures in both the clinical and basic hearing sciences.

Acknowledgments

This work was supported by grants from the Public Health Service (DC 00613, DC03114).

—Brenda L. Lonsbury-Martin

References


Further Readings


Otoacoustic Emissions in Children

Unidentified hearing loss in infants and young children can lead to delays in speech and language acquisition (Yoshinaga-Itano et al., 1998; Moeller, 2000). Identification of hearing loss presents an additional challenge because these patients may be unable to provide voluntary responses to sound. Otoacoustic emissions (OAEs) are an effective means to identify hearing loss in young children because they are related to the integrity of the peripheral auditory system and do not require voluntary responses from the patient.

OAEs are by-products of normal, nonlinear cochlear function, the source of which is the outer hair cell (OHC) system. They may be evoked by single tones (stimulus frequency, SFOAEs), pairs of tones (distortion product, DPOAEs), or transient stimuli (transient evoked, TEOAEs), and take from several seconds to several minutes to measure. Although all OAEs do not require a behavioral response, only TEOAEs and DPOAEs have been used widely to identify hearing loss. Since OHC damage results in hearing loss, OAEs, which are generated by OHCs, should be present when cochlear function is normal and reduced or absent when it is not. These facts have led to the application of OAE measurements in efforts to describe auditory function in humans, especially infants and young children. Unfortunately, OAE response properties from normal and impaired ears are not completely distinguishable; thus, diagnostic errors are inevitable. Below, we provide brief descriptions of TEOAEs and DPOAEs in normal-hearing infants and young children, followed by a description of these two OAEs in patients with hearing loss. Robinette and Glattke (2002) and Hall (2000) provide more background information and extensive references on OAEs. Norton, Gorga, et al. (2000) and Gorga, Norton, et al. (2000) provide comprehensive descriptions of OAEs in the perinatal period.

Infants and young children produce larger OAEs than older children and adults (Prieve, Fitzgerald, and Schulte, 1997; Prieve, Fitzgerald, Schulte, and Kemp, 1997; see Widen and O’Grady, 2002, for a review). There are several explanations for this difference. Very young children have not been exposed to environmental factors that might result in OHC damage. Also, their middle ears transmit energy to and from the cochlea differently than adults (Keefe et al., 1993), which might alter OAE levels in the ear canal. In addition, infant ear canal resonances show greater level at high frequencies, compared to spectra measured in adult ear canals. If stimuli differ in adult and infant ear canals, then responses may differ as well. Finally, the space between the measuring microphone and the eardrum is smaller in infants than in adults. If equivalent OAEs were generated in the cochlea, that signal would be larger in the infant ear canal because it was recorded in a smaller space.

While differences in infant and adult OAEs exist, the larger question revolves around whether OAEs can be used to distinguish ears with hearing loss from those with normal hearing. This dichotomous decision is made whenever OAEs are used in screening programs, regardless of the target population. The following discussion describes work in this area.

The clinical value of OAE measurements was recognized starting with their discovery (Kemp, 1978). Several studies describe the accuracy with which OAEs identify auditory status (e.g., Martin et al., 1990; Prieve et al., 1993; Gorga et al., 1993a, 1993b, 1996, 1997, 1999, 2000; Glattke et al., 1995; Kim et al., 1996; Hussain et al., 1998; Dorn et al., 1999; Harrison and Norton, 1999; Norton, Widen, et al., 2000). In general, both TEOAEs and DPOAEs identify auditory status with greater accuracy for middle and high frequencies than for lower frequencies. This occurs because noise levels decrease as frequency increases during OAE measurements. The noise interfering with OAE measurements (1) is acoustical, (2) results mainly from patient breathing and/or movement, and (3) contains mostly lower frequency energy. Noise adds variability and reduces measurement reliability. Thus, OAE test performance depends heavily on the frequencies at which predictions about auditory status are being made, in large part because noise level depends on frequency.

Test performance also depends on stimulus level (Whitehead et al., 1995; Stover et al., 1996; Harrison and Norton, 1999). Moderate-level stimuli result in the fewest false positive and false negative errors. Lower or higher stimulus levels decrease one of these error rates at the expense of increasing the rate of the other. This occurs for simple reasons. If low stimulus levels are chosen, virtually every ear with hearing loss will fail the test, resulting in a false negative rate of zero. However, the number of ears with normal hearing not producing responses will also increase as stimulus level decreases, increasing the false positive rate. If high-level stimuli are used, the vast majority of ears with normal hearing will produce responses, resulting in a low false positive rate. Unfortunately, some ears with hearing loss, especially ears with mild or moderate losses, will produce a response to high-level stimulation, increasing the false negative rate. Moderate-level stimuli result in optimal combinations of false positive and false negative rates. Thus, primary levels of 50–65 dB SPL for DPOAE measurements or 80–85 dB pSPL for clicks during TEOAE measurements are recommended.

Figure 1 shows representative examples of DPOAE and TEOAE signal and noise levels for three hearing loss categories. In general, robust responses above the noise floor are observed when hearing is normal (top row). When borderline normal hearing or mild hearing loss exists (middle row), the response is either reduced in level or absent. In cases of moderate or greater hearing loss (bottom row), the response typically does not exceed the noise floor, even when the noise level is low. These examples are consistent with general response patterns in these hearing loss categories, but it is important to
remember that both measurements will produce some diagnostic errors.

One of these errors (false positives or false negatives) may be more important for certain clinical applications. For example, infants or young children who are brought to a speech and hearing clinic or an otolaryngology clinic because of concern about hearing loss are at higher risk than the general population. In this case, one might choose stimuli or criteria that provide higher sensitivities, despite the higher false positive rates, because missing a hearing loss is the greater concern. In contrast, one might choose stimuli or criteria that provide higher specificity, despite the lower sensitivity, when the target population includes only well babies without risk for hearing loss. In this group, the probability of hearing loss is so low that it may be more important to minimize false positive errors. It is impossible to recommend a single set of stimuli and/or criteria because individual clinics must decide which error is more important for their needs.

OAE test performance also depends on the audiometric criterion defining the border between normal hearing and hearing loss. Both TEOAEs and DPOAEs perform best when thresholds ≤ 20–30 dB HL are used to define normal hearing. There are data suggesting that TEOAEs are more sensitive than DPOAEs to mild hearing loss (for a review, see Harris and Probst, 1997). However, direct comparisons failed to reveal large differences in test performance between TEOAEs and DPOAEs when audiometric criterion was varied (Gorga et al., 1993b; Norton, Widen, et al., 2000). Still, some differences across frequency have been observed. TEOAEs tend to perform better at detecting hearing loss for lower frequencies while DPOAEs tend to perform better at detecting high-frequency hearing loss (Gorga et al., 1993b; Kemp, 1997), because of how each measurement is made.

During TEOAE measurements, a fast Fourier transform (FFT) is performed on the ear canal waveform. The status of the cochlear region associated with specific frequencies is determined by examining the energy (or signal-to-noise ratio) at those frequencies. For example, one would conclude that the 1000-Hz region of the cochlea is functioning if energy is observed in the FFT at 1000 Hz. For DPOAE measurements, two tones are presented simultaneously (f1 and f2), interact at the cochlea place close to where f2 is represented, and produce distortion products (DP), the most prominent of which is the 2f1−f2 DP. The level of this component is then measured to determine if the cochlea is functioning at the point of its initial generation (f2). However, 2f1−f2 occurs at a frequency that is about one-half octave lower than f2. Thus, the measured response may occur in a region in which noise floors are less favorable, thus reducing measurement reliability. As a consequence, DPOAEs are less accurate than TEOAEs for lower frequencies.

Figure 1. OAE (circles) and noise (triangles) levels as a function of frequency, with DPOAE and TEOAE data shown in the left and right columns, respectively. Note that the y-axis is not the same for DPOAEs and TEOAEs. Representative examples are shown for normal hearing (top row), borderline normal hearing/mild hearing loss (middle row), and moderate to severe hearing loss (bottom row). See Gorga et al. (1997) for an explanation of the shaded areas in the DPOAE column. Although Hussain et al. (1998) developed a similar grid for TEOAE data, those data were collected with a different paradigm, and thus cannot be applied to the present set of data.
Acknowledgments

Much of the authors’ work summarized in this chapter was supported by grants from the NIH (NIDCD R01 DC2251 and DC R10 1958).

See also Otoacoustic Emissions.

—Michael P. Gorga, Stephen T. Neely, and Judith E. Widen

References


Otoacoustic Emissions in Children 517

References


Ototoxic Medications

Many medications are ototoxic, meaning that they adversely affect inner ear function. The toxicity can be divided into two broad categories, cochleotoxicity and vestibulotoxicity. Cochleotoxic medications affect hearing function and typically manifest with tinnitus (an abnormal noise in the ear), decreased hearing, or both. Most cochleotoxins affect hearing at the highest frequencies first, reflecting damage to hair cells of the cochlea. Vestibulotoxic medications affect the part of the inner ear that senses motion—the vestibular system. Vestibulotoxicity usually manifests with dizziness, unsteadiness, and when severe, oscillospsia. Oscillospsia denotes inability of a person to see when the head is moving, although visual acuity may be normal when the head is still. The most common vestibulotoxins, the aminoglycoside antibiotics, primarily damage the vestibular hair cells. Ototoxicity is often irreversible, as humans lack the ability to regenerate hair cells.

Ototoxic medications can be broken down into several broad groups. Chemotherapeutic agents are often cochleotoxic. Many antibiotics are ototoxic, and those in the aminoglycoside family are all ototoxic to some degree, some being primarily cochleotoxic and others primarily vestibulotoxic. Diuretics, mainly the loop diuretics, are often cochleotoxic. Similarly, quinine derivatives such as antimalarials are also commonly cochleotoxic. Many common medications in the non-steroidal anti-inflammatory group, such as aspirin, are cochleotoxic.

Chemotherapy Agents

Chemotherapeutic agents are drugs generally used to treat cancer. Actinomycin, bleomycin, cisplatin, carboplatin, nitrogen mustard, and vincristine have all been reported to be ototoxic. Their ototoxicity is generally a direct result of their toxicity to other cells. While cochleotoxic, these medications are rarely encountered as a source of vestibular dysfunction.

Cisplatin is currently the most widely used anticancer drug, and unfortunately, it is cochleotoxic. The toxicity of cisplatin is synergistic with that of gentamicin (Riggs et al., 1996), and high doses of cisplatin have been reported to cause total deafness. In animals, cisplatin otoxicity is related to lipid peroxidation, and the use of antioxidant agents is protective (Rybak et al., 2000). Some chemotherapy medications also have central nervous system toxicity, which can be confused with vestibulotoxicity.

Antibiotics

A large number of antibiotics have been reported to be ototoxic in certain circumstances, including erythromycin, gentamicin, streptomycin, dihydrostreptomycin, tobramycin, netilmicin, amikacin, neomycin, kanamycin, etiomycin, vancomycin, and capreomycin. Antibiotics generally considered safe are members of the penicillin family, the cephalosporin family, and the macrolide family (except in situations where dosages are very high). Here we will discuss a few of the more common ototoxic antibiotics.

The aminoglycosides are a large family of antibiotics that are uniformly ototoxic. Streptomycin, the first clinically used aminoglycoside, is now used only in treating tuberculosis, because many gram-negative bacteria are resistant and because of substantial ototoxicity. Dihydrostreptomycin is no longer used in the United States, but streptomycin sulfate can still be obtained. Streptomycin is primarily a vestibulotoxin.

Neomycin, isolated in 1949, is now mainly used topically because of renal toxicity and cochleotoxicity. Hearing ototoxicity from oral absorption of neomycin has been reported (Rappaport et al., 1986) and there may also be toxicity from eardrops in patients with perforated eardrums.

Vancomycin, developed in 1957, has been replaced by newer aminoglycosides such as gentamicin, tobramycin, netilmicin, and amikacin. It is not thought to be as ototoxic as neomycin.

Gentamicin is presently the biggest problem antibiotic with respect to ototoxicity, as most of the other ototoxic antibiotics have been replaced. Gentamicin was released for clinical use in the early 1960s (Matz, 1993). Netilmicin has equivalent ototoxicity to gentamicin (Tange et al., 1995). Hearing toxicity generally involves the high frequencies first, but it is rarely severe. Vestibulotoxicity, rather than hearing toxicity, is the major problem resulting from gentamicin use. Certain persons with mitochondrial deletions in the 12S subunit are much more susceptible to gentamicin than the general population (Fischel-Ghodsian et al., 1997). The prevalence of this mutation is not clear, but 1% of the population is a reasonable estimate, based on available data. Gentamicin accumulates in the inner ear with repeated dosing, and for this reason, most cases of toxicity are associated with durations of administration of 2 weeks or more.

Vancomycin, although not an aminoglycoside, does have minor ototoxicity. However, it is often combined with aminoglycosides, and in this situation it potentiates the ototoxicity of gentamicin (Brummert et al., 1990) as well as (probably) other aminoglycosides such as tobramycin. Vancomycin by itself, in appropriate doses, is not ototoxic (Gendeh et al., 1998). Occasional persons do appear idiosyncratically to have substantial vestibular toxicity from vancomycin. The reason why occasional persons are more sensitive is not clear but might resemble the situation with gentamicin, where there is a susceptibility mutation (Fischel-Ghodsian et al., 1997).

Eardrops may contain antibiotics, some of which can be ototoxic when administered to persons with perforated eardrums. Cortisporin otic solution appears to be the most ototoxic to the cochlea of guinea pigs. Ofloxacin eardrops have negligible toxicity (Barlow et al., 1995). Neomycin-containing eardrops have been reported to contribute to hearing loss (Podoshin, Fradis, and Ben David, 1989) in a relatively small way, but a definitive assessment of risk has not yet been made. The
vestibulotoxicity of eardrops has so far not been studied, although case reports suggest that gentamicin-containing drops are toxic (Marais and Rutka, 1998; Bath et al., 1999).

There are several known interactions between antibiotics as well as with other agents. Vancomycin combined with gentamicin causes more vestibulotoxicity than either one alone. Loop diuretics potentiate aminoglycoside toxicity. Noise also may potentiate aminoglycoside ototoxicity.

Delayed ototoxicity, meaning essentially toxicity that continues for several months after the drug has been stopped, occurs because the aminoglycosides are retained within the inner ear much longer than in the blood. Gentamicin has been reported to persist for more than 6 months in animals (Dulon et al., 1993). Neomycin, streptomycin, and kanamycin are also known to be eliminated from the inner ear slowly (Thomas, Marion, and Hinojosa, 1992).

**Ototoxic Diuretics**

Loop diuretics are well-known cochleotoxins. Examples include furosemide and ethacrinic acid (Rybak, 1993). Diuretics generally considered safe include chlorothiazide.

Loop diuretics are rarely a source of vestibulotoxicity. They are possibly a source of hearing disturbance. They may be synergistic with aminoglycoside ototoxins such as gentamicin, neomycin, streptomycin, and kanamycin. It seems prudent to attempt to avoid exposure to these agents if hearing is impaired.

**Quinine Derivatives**

Numerous quinine derivatives, including quinidex, atabrine, plaquenil, quinine sulfate, mefloquine (Lariam), and chloroquine, have reported ototoxicity (Jung et al., 1993). The toxicity is primarily cochleotoxic and is generally confined to tinnitus, but it can cause a syndrome that includes tinnitus, sensorineural hearing loss, and vertigo. Some quinine derivatives taken for malaria prevention can occasionally cause significant and long-lasting tinnitus. Recent studies suggest that quinine impairs outer hair cell motility.

**Aspirin, NSAIDs, and Other Analgesics**

Aspirin and other NSAIDs are commonly used, and apparently only toxic to hearing (Jung et al., 1993). These include ibuprofen, naproxen, piroxicam, diflunisal, indomethacin, etodolac, nabumetone, ketorolac tromethamine, diclofenac sodium, and the salicylates, aspirin and salsalate.

Rarely, hearing loss is reported from other types of analgesics, for example, hydrocodone/acetaminophen combination (Friedman et al., 2000; Oh, Ishiyama, and Baloh, 2000).

Permanent hearing disturbances are possible but rare. They are most commonly seen in individuals who take aspirin in large doses for long periods, such as for the treatment of arthritis. Occasionally persons with Ménière’s syndrome will develop a hearing disturbance from a small amount of an NSAID.

**Compounding Factors**

Noise exposure is the most common source of hearing loss. Industrial exposure characteristically causes a “noise notch,” with the hearing loss at mid- to high frequencies bilaterally. Guns and other unilateral sources of noise can cause more circumscribed lesions. Noise can be a cofactor in medication-induced ototoxicity. Those who have hearing loss from an ototoxic antibiotic, for example, may be at much greater risk from noise (Aran, 1995).

**Protection from Ototoxins**

Little is known about protection from ototoxicity. Antioxidants protect partially from noise or toxins in several animal models (Rybak, Whitworth, and Somani, 1999). In theory, prevention of reactive oxygen species, neutralization of toxic products, and blockage of the apoptosis pathway might provide protection from oxidative stress, which is a common final pathway for ototoxicity. Toxic waste products can be neutralized with glutathione and derivatives (Rybak et al., 2000). Apoptosis can be blocked using capsase inhibitors. At this writing, all of these approaches are investigational and are not being used clinically. Most also require delivery systems that go directly into the inner ear, and are therefore impractical for clinical use. For cochleotoxicity, noise avoidance is likely helpful, but even here the story is complicated. Paradoxically, moderate amounts of noise may protect from extreme amounts of noise. Nevertheless, it seems prudent to avoid excessive noise exposure, particularly in situations where there has been a recent exposure to an ototoxic. Because aminoglycosides may persist in the inner ear for more than 6 months, practically this advice implies long-term noise avoidance.

**Treatment of Ototoxicity**

Because most cochleotoxicity is caused by damage to hair cells, and because once dead, hair cells do not regenerate in humans, treatment of completed ototoxicity, whether it be cochlear or vestibular, is limited to substitution of other inputs, procedures that recalibrate remaining function, and behavioral adaptations. For cochleotoxicity, the approach is largely amplification. Hearing aids and related devices (assistive devices such as telephone amplifiers) can be helpful in those whose hearing loss is subtotal. When hearing loss is complete, cochlear implants may be offered.

For vestibulotoxicity, a rehabilitation approach is often very helpful. The goal of vestibular rehabilitation is to reduce symptoms of dizziness, oscillopsia, and unsteadiness. Patients are instructed in and perform a daily exercise routine designed to recalibrate remaining vestibular input and to substitute other senses such as vision and neck proprioception.
With reduced vestibular function, the eye movement generated for a given head movement is too small, resulting in oscillopsia or blurred vision. To train the brain to generate an eye movement of equal amplitude to the head movement, gaze stabilization exercises are performed. Exercises consist of focusing on an object with continuous movements of the head for 1–2 minutes. Exercises can be made more difficult by increasing the complexity of the visual background behind the object of regard. The complexity of the exercises is progressed gradually as symptoms resolve.

For the balance deficits that are common in vestibular ototoxicity, patients are given static and dynamic exercises that require the control of balance with reduced sensory input, conflicting sensory input, reduced base of support, and during head movements. To help adapt to their vestibular loss, patients are educated in environmental and behavioral modifications to reduce the risk of falls.

—Timothy C. Hain and Janet Helminski

References


Pediatric Audiology: The Test Battery Approach

The auditory mechanism is a complex sensory system and as such requires a wide selection of specific procedures for assessing its functional integrity. In general, these procedures may be grouped according to the differential information they supply about peripheral auditory disorders (i.e., external and middle ear, cochlea, and cranial nerve VIII), central auditory dysfunction (i.e., neural pathways of the brainstem and auditory cortex), and pseudohypacusis (i.e., hearing loss of nonorganic origin). The results of a battery of procedures contribute specific information about the auditory processes that are normal as well as those that are abnormal.

In differential audiologic assessment, the audiologist seeks procedures that provide optimum information about which levels of the auditory system are disordered. Patients can—and often do—have coexisting disorders at several levels, with the most dominant problem masking clues to the presence of others. Because no single test can represent the integrity of the entire auditory system, the best overall measure is obtained by combining test results, whereby each test within the test battery evaluates some aspect of the auditory mechanism. Jaeschke, Guyatt, and Sackett (1994) propose that useful diagnostic tests distinguish among disorders or states that might otherwise be confused, add information be-
Some tests are designed specifically to assist in identifying the site of the lesion, while others are designed to determine the presence and nature of an auditory deficit. The diagnostic outcomes sought from the pediatric population vary little from adult counterparts. That is, audiological tests are selected to differentiate peripheral versus central hearing loss, conductive versus sensorineural hearing loss, cochlear versus neural site of lesion, and varying middle ear conditions. The audiologist neither expects nor gets complete agreement on all the different tests performed. Age, physical and cognitive/intellectual conditions, individual variability, and peculiarities of different audiological conditions can result in paradoxical outcomes and may affect the consistency of audiological findings. Thus, an extensive battery of audiological tests is intended to provide a profile of data that may be compared with findings obtained in individuals with previously documented auditory conditions. Evidence for a specific interpretation exists when the profile of results is consistent with expected findings.

The test battery approach in pediatric audiology is focused on confirming suspected hearing loss in infants referred from universal newborn hearing screening programs and the ongoing assessment of infants at risk for delayed-onset or progressive hearing loss (Joint Committee on Infant Hearing, 2000). Refinements in audiological tests (e.g., conditioned behavioral tests including visual reinforcement audiometry and conditioned play audiometry; acoustic immittance; auditory-evoked potentials), as well as the addition of new audiological tests (e.g., otoacoustic emissions), provide the audiologist with a sophisticated test battery from which to initiate clinical decisions (Folsom and Diefendorf, 1999).

When individual tests are combined into a test battery, results can be viewed from a holistic framework. In this approach, findings across tests are integrated to establish a working diagnosis that often goes beyond the sum of the individual parts. The use of a test battery offers several advantages, including (1) avoidance of overgeneralizing the results from a single test, (2) increasing the data set from which to draw conclusions, and (3) enhancing the confidence in a clinical decision as the number of test results consistent with a specific interpretation increases. Conversely, combining tests into a battery may not be advantageous, cost-effective, or time-efficient when the tests are highly correlated (that is, different tests testing for the same disorder). The more positive the test correlation, the less performance varies when tests are combined. When tests have high to maximum positive correlation, test battery performance cannot be better than the best single test in the battery; thus, there is no value in combining tests just to satisfy the faulty assumption that more tests are always better. Each test must be selected on the basis of the patient’s complaints, and associated with the highest hit rate and lowest false alarm rate for the suspected disorder.

When selecting tests as part of a test battery, it is essential to balance quality patient care with fiscal responsibility. Therefore, the general rule to apply in pediatric assessment when selecting appropriate audiological tests is not to administer a test unless its results provide new information for patient management. In fact, the real advantage of tests in a test battery comes when negative correlation is determined between tests, indicating that each test tends to identify different disorders.

Not only does the test battery delineate hearing loss, it also provides opportunities for making appropriate cross-checks. The cross-check principle in audiology, originally outlined by Jerger and Hayes (1976), undergirds the concept of a test battery approach so that a single test is not interpreted in isolation, but various tests act as a cross-check on the final outcome. The principle is that the results of a single test are never accepted as conclusive proof of the nature or site of auditory disorder without support from at least one additional independent test. That is, the error inherent in any test and in patient response behavior is recognized, and the probability of an incorrect diagnosis is minimized when the results of several tests lead to the same conclusion. Moreover, the test battery approach and cross-check principle provide a statistical advantage when compared with the utilization of but a single test. The multiplicity of judgments in the test battery renders the entire differential assessment more reliable and valid. Statistically, multiple judgments from nonduplicative, negatively correlated tests lend safety to the interpretation of raw data when compared to the outcome and potential for error from a single judgment. To implement the cross-check strategy successfully, clinicians need to recognize the importance of selecting tests based on the child’s physical status, developmental level, and test correlation.

A test battery is paramount when the clinician is evaluating children with multiple disabilities. These children exhibit diverse medical problems that can diminish the accuracy of behavioral and physiological hearing tests. Complicating factors may include but are not limited to severe neurological, motor, and sensory problems. These factors can adversely influence test results, in turn compromising the validity of a single test approach. In addition, the limitations of the tests themselves can impose barriers when evaluating children with complex problems. Test constraints (e.g., limitations of behavioral observation audiometry [BOA] in eliciting an observable response; impact of developmental age on visual reinforcement audiometry [VRA]; middle ear pathology compromising acoustic reflex measures; the impact of central nervous system damage on the auditory brainstem response [ABR]) frequently dictate what procedures are feasible for a child with special needs. No test or clinician is infallible, and mistakes made with infants and young children can have crucial implications for medical and educational management.

Gans and Gans (1993) tested children with special needs by a test battery made up of BOA, VRA, ABR, and the acoustic reflex. The primary goal of the study was to rule out bilateral hearing loss greater than a mild degree. Stringent criteria were established for ruling
out hearing loss with each of the tests to minimize the chances of missing a child (false negative error) with a moderate hearing loss or greater. The tests were performed in a serial manner until one test result ruled in essentially normal hearing. Once achieved, further testing was discontinued.

The results demonstrated that BOA passed approximately 35% of the children, VRA passed approximately 10%, acoustic reflex measurement passed approximately 22%, and ABR passed approximately 57%. Yet when conducted in a serial strategy (individual tests administered until one “normal” result was obtained), 80% of the children under study were determined to have hearing better than the cutoff criterion. Although ABR alone was better at predicting hearing sensitivity than the other tests, the total percentage score accomplished by a serial test battery was more than 20% greater than for ABR alone. Factor analysis failed to find a strong relationship among the tests and suggests that different factors caused changes in threshold estimation across different children. That is, successful outcomes were based on interactions between the individual’s disabilities and the individual tests. These factors included but were not limited to low chronological or developmental age (neurological, motor, skeletal, and respiratory abnormalities), medications, and conductive hearing loss. Certain factors will adversely influence the results of one test more than another. Therefore, reliance on a single test for a child with disabling conditions would give an erroneous clinical impression that a large proportion of these children sustain hearing loss.

An audiologist prone to using a single-test approach might be tempted to rely on ABR as the only test method. This approach certainly relies on the assumption that the test of choice is a valid hearing test for all individuals and is not susceptible to variables that could lead to errors in outcome. The important findings from the work of Gans and Gans (1993) provide evidence that challenges this assumption for children with special needs, and provides a strong rationale for using a battery of appropriately selected tests.

The selection of individual tests for use in a test battery must be supported by clinical and experimental evidence. If individual tests and their use in test batteries are not evidence based, are not cost-effective in outcomes, and do not positively impact patients, we diminish the quality of services provided.

—Allan O. Diefendorf and Michael K. Wynne

References


Further Readings


Physiological Bases of Hearing

Outer and Middle Ears

The cartilaginous pinna on the outside of the skull has a set of characteristic folds and curves, different for each individual, which set up a series of shadowings and reflections of the sound wave. The result is that the spectrum of the sound, as transmitted to the concha (the opening of the ear canal), is modified according to the direction and elevation of the sound’s source. Although the main cue for sound localization comes from comparing the relative intensities and times of arrival of the stimuli at the two ears, that comparison does not give us information on the elevation of the sound source, or whether the source is behind or in front of the head; that information is provided by the pinna. Moreover, the outer ear means that sound localization of a sort can be undertaken with only one ear.

The middle and inner ears are protected from the outside world by an ear canal, and the inner ear is further protected by a middle ear cavity. The closed cavity of the middle ear, however, is a common site for infection, in which pus and secretions in early stages, and the formation of fibrous tissue in later stages, reduce the efficiency of transmission of vibrations to the cochlea.

The relatively dense, incompressible cochlear fluids, enclosed in a bony canal, with their movement limited by the membranes of the inner ear, need a higher pressure of vibration for a certain amplitude of move-
ment than do sound waves in air. One job of the outer and middle ears is to transform the ratio (pressure/amplitude) of vibration from a low value suitable for the external air to a much higher value able to drive the cochlear fluids efficiently. This is undertaken by (1) acoustic resonances in the outer ear canal, (2) the footplate of the stapes in the oval window being much smaller than the tympanic membrane (so that forces from the sound vibration are concentrated in a small area), and (3) a lever action in the vibration both of the middle ear bones and of the tympanic membrane.

Vibration through the middle ear is affected by the middle ear muscles, the tensor tympani and the stapedius muscle, with contraction of the muscles reducing sound transmission. These muscles contract in response to self-produced activity such as vocalizations, but also in response to loud sounds, to give some partial protection of the inner ear against acoustic trauma.

The Cochlea

The cochlea performs a spectral analysis, sorting the incoming mechanical vibration into its different frequency components, and transduces the sound, turning the mechanical vibration into an electrical change that activates the fibers of the auditory nerve.

The cochlea, deep inside the temporal bone, has a spiral central cavity, which curves the long (35-mm) cochlear duct into a small space 10 mm across. The canal is divided by membranous partitions into three spaces, or scalae, which run the length of the canal (Fig. 1).

The mechanical vibration, transmitted to the cochlear fluids, causes a ripple-like motion (i.e., a traveling wave) in the membranes dividing the scalae and hence in the organ of Corti, causing deflection of the stereocilia or hairs on the hair cells (see Robles and Ruggero, 2001, for a review). Cyclical deflection of the stereocilia opens and closes the mechanotransducer channels, causing positive and negative potential changes within the hair cells. There are two types of hair cell (Fig. 2). In the outer hair cells, operation of the mechanotransducer channels with the resulting changes in electrical potential induce, in ways that are controversial, a mechanical response that enhances the initial mechanical vibration. The result is that the amplitude of the mechanical traveling wave grows exponentially as it travels along the cochlear duct away from its point of introduction. Because the dimensions and stiffness of the cochlear duct and membranes change along the duct, at some point along the duct, the ratio of mass of fluid that has to be moved by the introduced vibration to the stiffness of the membranes that assist in the moving becomes too great to permit the vibration to continue for that particular frequency of stimulation, and the wave dies out rapidly. For a single tone, the traveling wave therefore has a peak which is sharp and narrow, so that maximal stimulation of hair cells occurs along only a short region of the cochlear duct. The peak occurs near the base (i.e., oval window end) for high-frequency tones and near the apex for low-frequency tones; for a spectrally complex stimulus, the spatial pattern of vibration reflects the spectrum of the incoming sound. This is known as place coding of frequency. In addition, the time pattern of...
vibration at any point reflects the time pattern of the (spectrally filtered) acoustic stimulus.

The outer hair cells are particularly vulnerable components of the cochlea, being readily damaged by insults such as loud sound, anoxia, and many ototoxic drugs. They are also particularly vulnerable to degenerative changes. The result of such processes is that the traveling wave is reduced in amplitude, and the sharpness of its peak is reduced. The inner hair cells are still able to detect the vibrations, but the cochlea loses sensitivity (i.e., there is a hearing loss), and there is a degradation in its ability to perform the spectral analysis.

**Processing in the Auditory Nervous System**

The inner hair cells make synaptic connections with the afferent fibers of the auditory nerve, so that when the inner hair cells are stimulated in parallel with the outer hair cells, the auditory nerve fibers are activated, and both the spectral and temporal patterns of vibration are thereby signaled to the central nervous system. At low frequencies, the time pattern of neural firing can follow the time pattern of the vibration of the organ of Corti at the point of innervation. This is known as the temporal coding of sound frequency. However, at higher frequencies, the fluctuations cannot follow each cycle of the waveform, and therefore the stimulus can be signaled only by a change in the mean firing rate. At low frequencies, therefore (below about 300 Hz), both place coding and time coding can contribute; at high frequencies (above a few kHz), only place coding is operative. However, temporal fluctuations in the envelope of the sound waveform, if below a few hundred Hz, are still represented in the overall pattern of the firing.

Processing in the auditory brainstem extracts and enhances three features of the auditory stimulus. Spectral contrast in the sound stimulus is enhanced, temporal transitions are emphasized, and information on the sound locus is extracted.

The enhancement of spectral contrast depends on the spatial pattern of activity in the auditory nerve, i.e., on the place coding of sound frequency, which is then emphasized by neural lateral inhibition between adjacent cell groups in the central auditory system, starting at the cochlear nucleus. Auditory neural responses therefore become dominated by representations of the most intense spectral peaks (such as vowel formants in the case of speech), while responses to a lower level or a more spectrally uniform background stimuli are reduced. In some way that is not understood, the two mechanisms of frequency representation, namely, place coding and time coding, are neurally integrated into a single percept.

Temporal information in the sound waveform is also emphasized; the enhancement of temporal transitions in the neural responses means that the response to a fluctuating stimulus consists substantially of bursts of neural activity when the stimulus intensity increases, no activity when the intensity decreases, and very low levels of activity during steady portions of the stimulus.

The spatial location of the sound source is primarily analyzed by comparing the sounds arriving at the two ears. A stimulus on, say, the left will strike the left ear first, and will also be more intense in the left ear. Both of these cues are extracted by the nervous system to give an indication of direction, with a neural representation in the central nervous system such that neurons on one side of the body are driven most strongly by sounds originating on the opposite side. The time cues are extracted by comparing the time of arrival of the nerve impulses from the two ears at the medial superior olivary nucleus in the brainstem, while intensity differences are detected primarily in the lateral superior olivary nucleus. The information as to the location of the sound source is integrated in the inferior colliculus with spectral and temporal pattern information that had been enhanced at the cochlear nucleus, before reaching the auditory cortex via the specific thalamic nucleus, the medial geniculate body.
**Cortical Analysis**

It is likely that there is a division of function between different parts of the auditory cortex, situated on the superior surface of the temporal lobe, buried in the lateral (or sylvian) fissure. The more dorsal areas are likely to represent location in auditory space, while the more ventral areas are involved in the analysis of complex stimuli, such as of speech (Rauschecker and Tian, 2000). Recording from individual cortical neurons suggests that speech sounds are likely to be analyzed and represented only over a whole population, that is, they are represented as a pattern of activity that is spread over a large number of neurons, where consideration of activity of the whole population is necessary for the accurate specification of the speech sound. Within the population, individual neurons or neural assemblies may be specialized for the detection of critical features, such as spectral peaks or rapid temporal transitions, such as are necessary for the specification of the sound (Wang, 2000). One area traditionally associated with speech is Wernicke’s area, which lies just posterior to the primary cortical area in the dominant (generally left) hemisphere. Lesions of Wernicke’s area result in defects in comprehension and word selection. Further forward, on the ventral frontal lobe, lesions of Broca’s area result in deficits in the production of speech. While the lesion data show that these areas are critical, functional magnetic resonance imaging shows that listening to speech activates a much wider range of areas (Fig. 3), with much more extensive surrounding temporal, angular, and frontal areas likely to be involved in both linguistic and semantic analysis (Binder et al., 1997). Although these imaging studies give information on the cortical areas activated and suggest the potential for conceptually splitting a task into its different functional components, they do not provide any information on the neural mechanisms underlying the functions, which remain elusive.

—James O. Pickles

**References**


**Further Readings**


**Pitch Perception**

The study of pitch dates back to at least the time of Pythagoras, who formulated the relationship between the length of a string and the pitch it would produce if it were strummed. The perception of pitch is the basis of musical melody and the voicing of speech; moreover, pitch is an attribute of the sound created by many objects in our world.

While the common definition of pitch has to do with a subjective attribute of sound and is scaled from low to high, pitch is closely related to frequency, a physical attribute of sound. The other physical attributes of sound are level, temporal structure, and complexity (Rossing, 1990). The study of pitch perception is often linked to the ability of the auditory system to process the frequency content of sound. The physical attributes of sound are derived from the fact that sound occurs when objects vibrate. The rate at which an object vibrates in an oscillatory manner is the frequency of the sound. If an object vibrates back and forth in a regular and repeatable manner 440 times in 1 s, it is said to have a frequency of 440 cycles per second (cps), which is indicated as 440 Hz (Hertz). If this vibrating object generated sound, the sound would have a 440-Hz frequency. If the vibrating object were a guitar string, a musician would perceive the vibrating string to have a pitch of 440 Hz. A higher rate of vibration would produce a higher pitch and a lower rate a lower pitch. Thus, for the

[Figure 3. Functional magnetic resonance image of the human cortex during a speech analysis task. Areas that are heavily activated are shown in white. The section is sagittal through the left hemisphere. (Modified with permission from Binder, J. R., et al. [1997]. Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*, 353–362.)]
vibrating guitar string, pitch is the subjective attribute of frequency. The relationship between the physical attributes of sound and pitch is not as simple as the guitar-string example suggests, and a few of the complexities will be described later in this article. That is, frequency and pitch are not synonymous.

Frequency can be measured using physical or objective means to determine the rate of vibration (Rossing, 1990). The measurement of pitch requires perceptual measurement techniques found in the research toolbox of the psychoacoustician or the musician. In many contexts, the pitch of a test sound is determined by comparing its perceived pitch with that of a standard sound. The standard sound is usually a sound with a very regular vibratory pattern, such as a tone or a periodic train of brief impulses (a click train). If the comparison sound and the test sound are judged to have the same perceived pitch, then the pitch of the test sound is the vibratory frequency of the comparison sound. So, in the example involving the guitar string, the vibrating string would be perceived as having the same pitch as a tone or a click train with a 440-Hz frequency.

**Musical Pitch and Pitch Scales**

In music, pitch can have two different meanings. When the frequency of the vibrating guitar string doubles (e.g., from 440 to 880 Hz), the frequency has increased by an octave. Musical sounds that are an octave apart in frequency are perceived as similar, as opposed to a less similar perception for sounds that are not separated by an octave. This similarity means that a melody played at one octave is very recognizable when it is played in other octaves. Thus, musicians will say that sounds that differ by an octave have the same pitch, even though they have different frequencies and may be judged to match a standard tone of a different frequency. Thus, pitch can have two meanings in music, one referring to octave relationships and one to the actual frequency of a standard tone of a different frequency. The standard tone, pitch chroma is sometimes used to refer to octave relationships involving pitch, while pitch height is used to refer to pitch as determined by vibratory frequency. For example, a 440-Hz tone and an 880-Hz tone have the same pitch chroma but different pitch heights.

The notes of the musical scale represent different ratio intervals within an octave. Thus, musical pitch can be measured in terms of musical notes (the 12 notes of A, B, C, D, E, F, G and the half-tones, or the sharps and flats). The octave can be divided into 1200 equal logarithmic units (ratio units) call cents. The 12 musical notes divide this 1200-cent octave into 12 equal units, so that each note is 100 cents (100 cents is sometimes referred to a semitone). Each note represents a different ratio from the beginning of the octave; e.g., the ratio between the notes C and G is one-fifth. The pitch of the 440-Hz vibrating guitar string is the note A in the middle range of the musical octaves, and the note C would have a pitch of 528 Hz in this same octave, or 264 Hz in the next lower octave. Therefore, another measure of pitch is the musical scale expressed either by musical notes or by cents.

The other measure of pitch is the mel scale (Stevens and Volkmann, 1940). The mel scale is an attempt to measure pitch on a scale from low to high, as implied by the definition of pitch given above. A sound with a tonal frequency of 1000 Hz has a pitch of 100 mels. A sound that is judged to have a pitch twice as high has a 200-mel pitch, while a pitch judged to be twice as low has a 50-mel pitch. Thus, the 1000-Hz tone serves as a referent against which the pitch of other sounds can be compared. A mel scale relates the perceived pitch of a sound in mels to another variable, most often frequency. So a sound with a pitch of 300 mels is one that is perceived as having a pitch that is three times higher than that of a 1000-Hz tone.

**Complex Pitch**

Sounds can have complex patterns of vibration and as such are made up of vibrations of many different frequencies. These complex sounds have a spectrum of many different frequencies. If certain frequencies are dominant in the sound’s spectrum, then the pitch of the sound may be perceived as representing the frequency of the dominant component in the spectrum. So the pitch of simple sounds with regular patterns of vibration or complex sounds with dominant frequencies in the sound spectrum is determined by these main frequencies. However, we can consider a complex sound that is the sum of the following frequencies: 300, 400, 500, 600, and 700 Hz. This complex sound will have a perceived pitch of 100 Hz, even though the sound contains no frequency at 100 Hz. Note that this complex sound consists of frequency components that are spaced 100 Hz apart and are all integer multiplies (harmonics) of 100 Hz. The missing 100-Hz component is the fundamental frequency of the complex, because it is the highest frequency for which all of the other components are integer multiplies (de Boer, 1976). The 100-Hz perceived pitch is that of this missing fundamental, and so this type of perceived pitch is often referred to as the “pitch of the missing fundamental” or “complex pitch.” Although this sound’s spectrum has no frequency at its pitch, the pattern of vibration oscillates at 100 Hz. Thus, it is possible that the pitch is not related as much to the frequency content of the sound as it is to the rate of oscillation. Or perhaps the 100-Hz spacing of the frequency components is the crucial dimension of the sound that determines the pitch of the missing fundamental. If the spectrum of this sound is changed to be 325, 425, 525, 625, and 725 Hz, the perceived pitch is now 104 Hz, which is neither the missing fundamental nor equal to the 100-Hz spacing of the frequency components. In addition, the sound does not have a dominant pattern of vibratory oscillations at 104 Hz (Patterson, 1973). Thus, neither simple frequency nor temporal properties of a sound can be used to predict complex pitch.

Complex pitches, such as the pitch of the missing fundamental, suggest that pitch is not simply related to
frequency. It is also the case that changes in other physical attributes of sound, such as sound level, can lead to a change in pitch, reinforcing the observation that pitch and frequency are not synonymous. Various theories of pitch processing have been proposed to account for the pitches of simple and complex sounds. These theories fall into three general categories (Moore, 1997). One set of theories, spectral theories, propose different means of using the sound’s frequency spectrum to determine pitch. Another group of theories, temporal theories, suggest that aspects of the temporal oscillation of the sound’s vibratory pattern are used by the auditory system to determine pitch. Then there are theories that use a combination of spectral and temporal aspects of a sound to predict its pitch (de Boer, 1976). Such theories have been proposed for nearly a hundred years, and no one theory has emerged that best accounts for all of the data related to pitch perception (Plomp, 1976).

Thus, whereas pitch is a major attribute of sound and is crucial to our ability to use sound to identify the objects in our world and to communicate, auditory science does not have a good explanation of how pitch is processed by the auditory system. It is clear that the way in which the auditory system processes both spectral and temporal information contributes to pitch processing, but the details of how these processes operate is still a mystery.

—William A. Yost

References


Further Readings


Presbyacusis

Presbyacusis is a general term to describe hearing loss in older persons or the hearing loss associated with aging. Hearing loss observed in older adults is a result of the combined effects of aging, long-term exposure to occupational and nonoccupational noise, the use of ototoxic drugs, diet, disease, and other factors. In this case, the term presbyacusis describes any hearing loss observed in an older person, regardless of cause. Alternatively, presbyacusis may refer specifically to the hearing loss that increases with chronological age and is related only to age-related deterioration in the auditory periphery and central nervous system (CNS).

Currently, about 75% of the 28 million hearing-impaired individuals in the United States are 55 years of age or older; the number of hearing-impaired individuals will increase as the population ages. Indeed, presbyacusis is the most prevalent of the chronic conditions of aging among men age 65 years and older, and the fifth most prevalent condition among older women, following arthritis, cardiovascular diseases, and visual impairments (National Center for Health Statistics, 1986). Age-related hearing loss in the United States has been well characterized by epidemiologic surveys such as the Framingham Heart Study (Moscicki et al., 1985), the Baltimore Longitudinal Study of Aging (Pearson et al., 1995), and the Epidemiology of Hearing Loss Study of the adult residents of Beaver Dam, Wisconsin (Cruickshanks et al., 1998). Figure 1 shows the systematic increase in thresholds with chronological age in Framingham subjects; thresholds at high frequencies are higher for male than for female subjects. In the Beaver Dam study, the prevalence of hearing loss was 45.9%, with hearing loss defined as average thresholds (0.5–4 kHz) greater than 25 dB in the worse ear. This is consistent with the prevalence reported in the Framingham study of 42%–47%. The prevalence of hearing loss in the Beaver Dam study varied greatly with sex and age, ranging from 10.2% for women 48–52 years of age to 96.6% for men 80–92 years of age. Another set of data, including hearing levels as a function of age, sex, and history of occupational noise exposure, is part of an international standard (ISO 1999, 1990). Database A from this standard may more closely represent aging effects on hearing, given that subjects were screened for occupational and other noise exposure history. Epidemiologic surveys also provide estimates of the genetic component of presbyacusis. Heritability coefficients suggest that as much as 55% of the variance in thresholds in older persons is genetically determined, are stronger in women than in men, and are comparable to those for hypertension and hyperlipidemia (Gates, Couropmitree, and Myers, 1999).

A remarkable and consistent age-related change in hearing occurs at frequencies above 8 kHz and begins as early as age 20–30 years (Stelmachowicz et al., 1989). Figure 2 shows thresholds in the conventional and extended high frequencies for younger and older adults,
grouped according to average thresholds from 1 to 4 kHz (Matthews et al., 1997). Thresholds are substantially elevated at frequencies above 8 kHz, even for individuals with nearly normal thresholds at lower frequencies.

To supplement pure-tone thresholds, the Hearing Handicap for the Elderly instrument (Ventry and Weinstein, 1982), a self-report questionnaire, has been used to compare older individuals’ assessment of their communication abilities with objective measures of hearing and with threshold-based estimates of hearing handicap, such as that recommended by the American Academy of Otolaryngology–Head and Neck Surgery (AAO, 1979; Matthews et al., 1990). Discrepancies between objective and subjective measures are common, and the variance in pure-tone thresholds within hearing handicap categories is large. Thus, whereas hearing loss is among the most prevalent chronic conditions of aging, the impact of hearing loss on communication abilities and daily activities of older adults varies greatly among individuals and is not accurately predicted from the pure-tone audiogram.

Studies of auditory behavior in older adults must separate age-related effects from those attributable simply to reduced audibility resulting from elevated thresholds. One experimental method to minimize the confound of reduced audibility is to include only older subjects whose pure-tone thresholds are equal to those of younger subjects. When changes in auditory behavior in older adults are observed that are not attributable to reduced audibility, they may be due to age-related changes in the auditory periphery, which provides an impoverished input to a normal auditory CNS, or to the combined effects of an aging periphery and an aging CNS. For many behavioral measures, it may not be possible to differentiate between these outcomes. This is particularly the case for tasks that require comparisons of temporal information across intervals of time or that assess binaural processing.

Other than effects related to their hearing loss, older adults probably do not have increased problems in speech understanding relative to younger individuals, as measured conventionally (i.e., monaurally, under earphones, with highly redundant signals). Indeed, some studies suggest that 70%–95% of the variance in monaural speech recognition scores may be accounted for by the variance in speech audibility (Humes, Christopherson, and Cokely, 1992). Figure 3 shows scores on several speech recognition tests for three age groups with nearly identical (within ~3 dB) mean thresholds from 0.25 to 8 kHz (Dubno et al., 1997). With hearing loss held constant across age group, speech recognition scores also remained constant. Nevertheless, age-related differences in speech recognition in noise may become
Evidence of age-related changes in the auditory system is revealed in the physiological properties of aging humans and animals. Older gerbils raised in quiet have elevated thresholds of the compound action potential (CAP) of the auditory nerve and shallow slopes of CAP input-output functions (Hellstrom and Schmiedt, 1990). These characteristics are also reflected in higher auditory brainstem response (ABR) thresholds and shallower slopes of ABR amplitude-intensity functions relative to young gerbils (Boettcher, Mills, and Norton, 1993). Similar findings have been observed in older humans. Although these potentials produced by short-duration signals are reduced in amplitude with age, amplitudes of potentials arising from higher CNS centers in response to long-duration signals, such as steady-state potentials and N100–P200, may be unaffected or even increase with age. Abnormal recovery from adaptation or forward masking and abnormal gap detection at the level of the brainstem have also been observed in older animals and humans (Walton, Orlando, and Burkard, 1999). In aging gerbils, the 80–90 mV dc resting potential in the scala media of the cochlea, known as the endocochlear potential (EP), is reduced substantially. In contrast to these changes, nonlinear phenomena remain relatively intact. For example, transient otoacoustic emissions, reflecting the functioning of outer hair cells, are present in about 90% of older humans with normal hearing, but amplitudes are reduced in a manner that is not predictable by either age or pure-tone thresholds. In aging gerbils, distortion product otoacoustic emissions are present and robust, but somewhat reduced in amplitude (Boettcher, Gratton, and Schmiedt, 1995). Two-tone rate suppression is observed in older gerbils, with age-related threshold shifts (Schmiedt, Mills, and Adams, 1990), although older humans may have reduced suppression measured psychophysically (Dubno and Ahlstrom, 2001).

The pathologic anatomy underlying these physiologic changes was most extensively described through studies of human temporal bones by Schuknecht (1974). Initially, four categories of presbyacusis were identified, including sensory (degeneration of sensory cells), neural (largely loss of spiral ganglion cells), metabolic (degeneration of the lateral wall and stria vascularis, with reduction in the protein Na,K-ATPase), and mechanical (aging of sound-conducting structures of the inner ear). Categories of presbyacusis were revised by Schuknecht and Gacek (1993) wherein atrophy of the stria vascularis was designated as the “predominant lesion” of the aging ear, neuronal loss was “constant and predictable,” mechanical loss remained theoretical, and sensory presbyacusis was the “least important type of loss.” These histopathological findings are consistent with the physiological evidence described above, such as reduced CAP amplitudes and reduced EP but robust otoacoustic emissions. Thus, most age-related changes in hearing can be accounted for by changes observed in the auditory periphery. Nevertheless, there are many age-related anatomical, neurochemical, and neurophysiological changes in the CNS. One prominent neurochemical change is a loss of gamma-aminobutyric acid, which may affect the balance of inhibitory and excitatory
neurotransmission. The effects of these and other CNS changes on age-related hearing loss may be substantial but remain largely unknown.

Acknowledgments

This work was supported (in part) by grants P50 DC00422 and R01 DC00184 from the National Institute on Deafness and Other Communication Disorders (National Institutes of Health).

—Judy R. Dubno and John H. Mills

References


Further Readings


Pseudohypacusis

Pseudohypacusis means, literally, a false elevation of thresholds. In pseudohypacusis, intratest and intertest audiometric inconsistencies cannot be explained by medical examinations or a known organic condition (Ventry and Chaiklin, 1965). Authors of literature in this area have called this condition exaggerated hearing loss, nonorganic hearing loss, or functional hearing loss. Exaggerated hearing loss implies intent, but some forms of pseudohypacusis may have subconscious origins (Wolf et al., 1993). The intent of the listener cannot be determined with audiometric measures. Nonorganic hearing loss implies that there is no physical basis for the hearing loss; however, many adults have a false elevation of thresholds added to an existing loss. Some even present with pseudohypacusis and an ear-related medical problem requiring immediate attention (Qui et al., 1998). Functional hearing loss is the only synonym among these terms.

Monetary or psychological gain motivates most pseudohypacusics. Audiologists should be alert for this condition if the referral source is a lawyer, as in a medicolegal case, or an organization documenting hearing for compensation purposes; however, there are cases where the referral source does not provide any warning. There are even cases of persons with normal auditory systems presenting with longstanding false losses that have been misdiagnosed and the individuals inappropriately fitted with hearing aids.

Estimates of the prevalence rate for pseudohypacusis are between 2% and 5%, with higher rates observed in some special populations, such as the military and industrial workers (Rintelmann and Schwann, 1999). Pseudohypacusics show falsely elevated thresholds in one or both ears, the degree of loss ranges from mild to profound, and the type of loss can be sensorineural or mixed (Qui et al., 1998).

Pseudohypacusics usually adopt an internal loudness yardstick that corresponds to the amount of their “hearing loss” (Vaubel, 1976; Gelfand and Silman, 1985). External sounds are compared with this internal yardstick, and pseudohypacusics only respond behaviorally to sounds that exceed this internal value. This is important to know, because modifications to hearing tests that affect loudness perception often have little or no effect on thresholds of audibility in cooperative adults. Many behavioral tests for pseudohypacusis that are used by audiologists are designed to disrupt loudness judgments.
The responsibility of the audiologist in the assessment of persons presenting with pseudohypacusis is to document intertest and intratest inconsistencies and to quantify true thresholds as a function of frequency. Many methods exist for documenting inconsistent results, but few measures exist for quantifying accurately true behavioral thresholds.

When evaluated for intertest and intratest inconsistencies, the basic battery of audiologic tests, administered to nearly everyone entering the clinic, are the ones that will likely identify persons presenting with this condition, given that many pseudohypacusics have nothing in their history that might raise suspicion. The basic battery of tests that is used in the assessment of pseudohypacusis differs somewhat across clinics, but this group of tests often includes pure-tone thresholds, spondee thresholds, and immittance (see *Tympanometry*).

Pure-tone threshold assessment provides several methods for identifying pseudohypacusis. Most clinicians routinely retest the threshold for a 1000 Hz tone as a reliability check. Thresholds on retest are usually within 10 dB of the first test in cooperative persons, whereas pseudohypacusics often show larger threshold differences (Ventry and Chaiklin, 1965). Although this method is not particularly sensitive or specific for the identification of pseudohypacusis, deviations greater than 10 dB can provide a warning to the clinician. In addition to poor reliability, pseudohypacusics do not demonstrate false positive responses (a response when a tone is not presented) (Ventry and Chaiklin, 1965). By contrast, several false positive responses in a single testing session are quite common in persons with tinnitus (Mineau and Schlauch, 1997). Unfortunately, audiometric configuration is not a reliable diagnostic tool for pseudohypacusis. Ventry and Chaiklin (1965), Kane (1985). The absence of shadow responses in asymmetrical losses is a reliable sign of pseudohypacusis (Rintelmann and Schwann, 1999). Shadow responses are thresholds based on the response of the non-test ear when sound is presented to the poorer, test ear. They reflect a limitation in the ability to isolate the two ears during a hearing test when masking noise is not presented to the non-test ear.

Spondee thresholds, a speech threshold for two-syllable words with equal stress on both syllables, are a quick measure that, when combined with pure-tone thresholds, provide one of the most effective tests for identification of pseudohypacusis. Spondee thresholds in cooperative adults usually fall within 10 dB of the average threshold for 500 Hz and 1000 Hz pure tones (PTA) (Carhart and Porter, 1971). Pseudohypacusics usually show larger differences, with the spondee threshold being lower (better) than the PTA (Carhart, 1952). In some instances, this difference may reflect the naiveté of the listener (Frank, 1976), as is often the case in children. In other words, the listener may feign a loss for tones and not understand that speech thresholds are quantifiable, too. However, this finding in most instances is a result of the loudness of speech and tones growing at different rates. Consistent with loudness-related issues, this test is most effective when spondee thresholds are measured using an ascending approach (beginning at a low level) and pure-tone thresholds are measured using a descending approach (beginning at a high level) (Schlauch et al., 1996). This procedure identified 100% of pseudohypacusics, with no incorrect identifications of cooperative test subjects with hearing loss. A more conventional procedure that measured pure-tone thresholds with an ascending approach identified only about 60% of persons with pseudohypacusis.

Immittance measures are also routine tests that can aid in the documentation of persons presenting with pseudohypacusis. Tympanometry and acoustic reflex thresholds are sensitive measures of middle ear status and provide some indication of the integrity of the auditory system up to the superior olivary complex. For severe sensorineural losses, acoustic reflex thresholds are generally 10 dB or more above a person’s behavioral threshold (Gelfand, 1994). Thresholds obtained at lower levels suggest pseudohypacusis. Gelfand (1994) has published normative values of reflex thresholds for different degrees of loss.

Numerous special tests were developed for assessing pseudohypacusis during the period immediately following World War II. Many of the tests developed during this time are confrontational, time-consuming, and ineffective when a clinical decision theory analysis is done. An exception to this criticism is a computer implementation of a simple modification to Bekesy audiometry or automated audiometry (Chaiklin, 1990), which still has application when testing the hearing of large groups of persons with a limited number of testers.

The Stenger test is a special test that is quick to administer and, unlike most tests, has the capability of quantifying the actual thresholds of persons presenting with unilateral pseudohypacusis (Kinstler, Phelan, and Lavender, 1972). This test makes use of the finding that when the same sound is presented simultaneously to both ears, the listener only hears the sound in the ear with the loudest percept. This test can be performed with tones or speech, but to be effective, the asymmetry between ears should be 40 dB or more. Manipulation of the sound levels in each ear is effective in identifying pseudohypacusis and quantifying actual behavioral thresholds (Rintelmann and Schwann, 1999).

Otoacoustic emissions (OAEs) (Musiek, Bornstein, and Rintelmann, 1995) and auditory-evoked potentials (Saunders and Lazenby, 1983; Bars et al., 1994; Musiek, Bornstein, and Rintelmann, 1995) are useful physiological measures for evaluating persons with pseudohypacusis. These special tests assess structures in the auditory pathways. OAEs assess the outer hair cells in the cochlea and the conductive pathway leading to the cochlea; these emissions are often absent even with mild hearing losses. However, auditory neuropathy cases, although rare, show that persons can have essentially normal OAEs and a severe hearing loss (Sininger et al., 1995). This
possibility and the fact that many adults exaggerate existing losses make interpretation of OAEs in isolation somewhat ambiguous when evaluating pseudohypacusis. Estimation of the auditory brainstem response (ABR) threshold, a type of evoked potential, has been advocated as a useful measure for compensation cases (Bars et al., 1994), but this threshold assessment tool, while accurate in many situations, can yield misleading results in certain hearing-loss configurations (e.g., a rising audiogram) (Glattke, 1993; Bars et al., 1994). ABR threshold, like OAEs, does not assess the entire auditory pathway as do behavioral measures. OAEs (Gorga et al., 1993) and ABR (Hall, 1992) are also ineffective tools for assessing low frequencies, a critical region for consideration in compensation cases. Other evoked potentials, such as middle latency responses and the slow cortical potential, may hold promise, but like the ABR test, they are expensive to administer (Hyde et al., 1986; Musiek, Bornstein, and Rintelmann, 1995).

The ABR threshold and OAEs are important for the complete documentation of intractable cases or persons presenting with severe to profound bilateral losses, but most cases of pseudohypacusis are resolved by a combination of readministering patient instructions, informing the patient that there are inconsistent responses, and making multiple measurements of the audiogram using an ascending approach. The validity of remeasured pure-tone thresholds is evaluated using the PTA-spondee threshold-screening test described earlier. Patients whose thresholds are not resolved using this approach are scheduled for additional testing. Persons with obvious psychological problems are referred for counseling.

See also CLINICAL DECISION ANALYSIS;OTOACOUSTIC EMISSIONS;PURE-TONE THRESHOLD ASSESSMENT;SUPRATHRESHOLD SPEECH RECOGNITION;TINNITUS.

—Robert S. Schlauch

References


Further Readings


**Pure-Tone Threshold Assessment**

*Audiometry* is the measurement of hearing. Clinical hearing tests are designed to evaluate two basic aspects of audition: sensitivity and recognition (or discrimination). Hearing sensitivity measures are estimates of the lowest level at which a person can just detect the presence of a test signal (Ward, 1964). Measures that require an identification response or judgments of sound differences are tests of auditory acuity, recognition, or discrimination. These tests provide information about a listener’s ability to recognize, discriminate, or understand acoustic signals, such as speech, and usually are conducted at moderate or higher signal levels. Tests of word recognition are clinical examples of such tests (see *SUPRATHRESHOLD SPEECH RECOGNITION*).

Hearing tests are performed for two primary purposes. One purpose is to identify hearing problems that may be caused by ear disease or damage to auditory structures. In some cases a hearing loss may indicate a medical problem, such as an ear infection. Because medical treatment of ear diseases is successful more often in early stages of the disease process, it is critical that such problems be detected and treated as soon as possible. A second purpose for hearing tests is to obtain information important for rehabilitation planning. In those cases of hearing loss for which medical treatment is not an appropriate alternative, it is important that non-medical rehabilitative measures be considered based on the communication needs of the affected individual. Information obtained from the hearing evaluation, for example, is required to make decisions about the need for personal amplification (such as a hearing aid) or for other auditory rehabilitation services.

A basic requirement for administration of hearing tests is an acoustic system that enables control of the signals presented to the listener. An *audiometer* is an electronic instrument used to present controlled acoustic signals to a listener in order to test auditory function. In conventional pure-tone audiometry, the audiometer provides for the presentation of tones ranging in frequency from 125 through 8000 hertz (Hz). A hearing level (HL) dial allows the tester to control the level (in decibels, or dB) of a tone being presented to the listener. The HL control is graduated in steps of 10 and 5 dB (and sometimes smaller) and typically is adjustable over a range of 120 dB. When the HL dial is set to 0 dB, the output level of the audiometer corresponds to an average normal HL at that specific frequency. This is referred to as audiometric zero (American National Standards Institute [ANSI], 1996). This instrumental convenience accounts for the differences in absolute hearing sensitivity (in dB sound pressure level, or SPL) across frequency in persons with normal hearing. Other controls on the audiometer enable the tester to route the test signal to various receivers or transducers used to present the tones to a listener. The two most common receivers used in pure-tone audiometry are a set of earphones and a bone conduction vibrator. Earphones provide for airborne acoustic signals; a bone conduction vibrator is used to transmit vibratory energy through the skull to the inner ear (cochlea). Diagnostic audiometers also provide masking signals (typically a noise) for presentation to the nontest ear during specific audiometric tests. A masking noise is necessary for bone conduction audiometry and for cases of substantial unilateral hearing loss in which the test signal may be intense enough to be heard in the nontest ear. To rule out this possibility, the masking noise is introduced in the nontest ear. The noise elevates the threshold in that ear (masks it) and eliminates it from
the test situation, to ensure that subject responses are only for the test ear.

Pure-tone audiometry consists of threshold measures for air and bone conduction signals in each ear (separately). A pure-tone threshold is the lowest level at which a person can just detect the presence of a tone. The threshold measure is statistical in nature; it is a level at which a listener responds to a criterion percentage of the signals presented. Clinically, threshold is usually defined as the lowest signal level at which the listener just detects 50% of the tones presented. Audiometric thresholds are influenced by a number of variables, including the instructions to the listener, the positioning of the earphone (or other transducer) on the head, and the psychophysical threshold measurement technique used (Dancer and Ventry, 1976; Yantis, 1994). In addition, individuals may demonstrate threshold variability related to factors such as motivation, the nature of the ear disorder, the patient’s ability to comply with the test situation, and ongoing physiological changes inherent to the auditory system (Wilber, 1999).

Threshold measures obtained from pure-tone audiometry are conventionally plotted on a graph called an audiogram (Fig. 1). The audiogram enables the tester to quickly see the extent to which thresholds for a listener deviate from normal. In Figure 1, HL (dB) is plotted on the linear vertical axis and signal frequency (Hz) is indicated on the logarithmic horizontal axis. The format of the audiogram is such that one octave along the frequency axis corresponds in dimensional scale to 20 dB on the HL axis (ANSI, 1996). Recommended symbols for audiograms are also included in the legend of Figure 1 (American Speech-Language-Hearing Association [ASHA], 1990). Note that separate symbols are used to indicate bone conduction thresholds and measures made with a masking noise in the non-test ear (masked thresholds). Symbols used to record thresholds on the audiogram also may be color coded, with red indicating measures for the right ear and blue indicating results for the left ear.

Specific procedures have been developed for pure-tone audiometry so that thresholds are obtained in a manner that minimizes test variability and are repeatable over time and from clinic to clinic. These procedures are based on research findings for persons with normal hearing and persons with hearing impairment, and include specifications on test frequencies, duration of test tones, step size changes in signal level, and other procedural variables (ASHA, 1978, 1997). The basic test protocol involves initially familiarizing the listener with the tones that will be heard and then determining thresholds for the tones. A bracketing technique is used whereby the tone level at a specific frequency is varied up and down and the listener indicates whether the tone is audible at each level. The tester determines the average hearing level at which the tone was heard approximately half the time over a series of presentations. This level represents the pure-tone threshold at a specific frequency and is recorded on the audiogram. Thresholds are obtained separately for air conduction using earphones and for bone conduction using a bone conduction vibrator. Air conduction tones produced by the earphone are directed down the ear canal, through the middle ear, and then to the cochlea. In bone conduction testing, however, a vibrator is used to transmit the signal through the bones of the skull to the cochlea. The vibrator is placed on the forehead or the mastoid portion of the temporal bone (behind the pinna), and thresholds are measured for the desired audiometric frequencies. All testing is performed in a sound-treated room that meets standards for the exclusion of ambient noise (ANSI, 1999).
Results of pure-tone audiometry, recorded on the audiogram, provide a description of both the degree and type of hearing loss. Because we are mainly interested in a person’s ability to hear everyday speech, the customary procedure for classifying degree of hearing loss involves a computation of the pure-tone average (PTA). A person’s PTA is her or his average pure-tone thresholds for the speech frequencies of 500, 1000, and 2000 Hz. Table 1 provides a degree classification of hearing loss based on average pure-tone thresholds in the better ear for these frequencies. In considering the purpose and results of audiometric tests, it is important to distinguish the terms hearing loss (or hearing impairment) and hearing handicap. 

Table 1. Degree Classifications of Hearing Loss (Handicap)

<table>
<thead>
<tr>
<th>Pure-Tone Average (dB HL) for 500, 1000, and 2000 Hz in the Better Ear</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤25</td>
<td>Not significant</td>
</tr>
<tr>
<td>26–40</td>
<td>Slight</td>
</tr>
<tr>
<td>41–55</td>
<td>Mild</td>
</tr>
<tr>
<td>56–70</td>
<td>Marked</td>
</tr>
<tr>
<td>71–90</td>
<td>Severe</td>
</tr>
<tr>
<td>&gt;90 dB</td>
<td>Extreme</td>
</tr>
</tbody>
</table>


Figure 2. Example audiograms for conductive (2a), sensorineural (2b), and mixed (2c) types of hearing loss. Further description is provided in the text.
handicap (or disability) (ASHA, 1997; ASHA/CED, 1998). The communicative handicap associated with a given amount of hearing loss in dB may differ considerably across individuals.

The classification of hearing loss according to type is based on comparisons of thresholds for air and bone conduction test tones. *Conductive* hearing losses are almost always due to abnormalities of the outer or middle ear. In such disorders, an interruption or blockage of sound conduction to the cochlea accounts for the hearing loss. Wax in the ear canal or breaks in the ossicular chain (middle ear bones) are examples of conditions that restrict the flow of sound from the outer ear and middle ear to the cochlea and result in a conductive hearing loss. The primary audiometric sign of a conductive hearing loss is an *air–bone gap*. An air–bone gap is present when hearing sensitivity by bone conduction is significantly better than by air conduction. This can occur in cases of outer and middle ear disorders because the path of sound transmission for bone conduction is primarily through the bones of the skull directly to the inner ear, essentially bypassing the affected outer and middle ear structures. An audiogram for a conductive hearing loss is shown in Figure 2A. Note that although the air conduction thresholds are elevated by 50–60 dB, bone conduction thresholds are within normal limits (0 dB HL). The primary effect of a conductive disorder is to reduce the level of sound reaching the inner ear. The impairment is primarily a loss in hearing sensitivity, not in speech understanding. If the sound level can be increased, the person will be able to hear and understand speech.

Hearing loss resulting from damage or disease to any portion of the inner ear or neural auditory pathways is classified as *sensorineural*. Because the problem lies in the inner ear or neural pathways (or both), there will be an equal hearing loss for both air- and bone-conducted signals. An audiogram for a patient with a sensorineural hearing loss is shown as Figure 2B. Notice that the bone conduction thresholds are the same as the air conduction thresholds—there is no air–bone gap. In contrast to conductive hearing loss, sensorineural hearing loss typically involves both a loss in hearing sensitivity and a reduced ability to understand speech. Even when speech is made louder, the person will still have some difficulty in understanding.

A *mixed* hearing loss is a combination of both conductive and sensorineural losses. Figure 2C is an example audiogram for a mixed hearing loss in both ears. Note that air conduction thresholds are poorer than normal, averaging about 60 dB HL, and bone conduction thresholds are also poorer than normal, averaging 35 dB HL. There is an air–bone gap of approximately 25 dB, suggestive of some conductive hearing loss. In addition, there is a loss by bone conduction, indicating some abnormality of the inner ear and/or auditory nerve.

The case of a mixed hearing loss underscores the fact that both sensorineural and conductive disorders can exist simultaneously in the same patient. A person with a significant sensorineural hearing loss, for example, can still experience an ear infection or other pathology that may result in a conductive component in addition to the sensorineural loss. Finally, it should be understood that many patients with significant disorders or pathologies of the auditory system may have no significant hearing loss (in dB). Patients with disorders of the central auditory nervous system, for example, may demonstrate normal pure-tone thresholds. Normal hearing sensitivity does not always indicate a normal auditory system (Wiley, 1988).

—Terry L. Wiley

**References**


**Further Readings**


The articulation index, or, as it is now known, the speech intelligibility index, was originally developed by early telephone engineer-scientists to describe and predict the quality of telephone circuits (Fletcher, 1921; Collard, 1930; French and Steinberg, 1947). Initially their motivation was to provide a method that would reduce the need for lengthy (and expensive) human articulation tests to evaluate the merit of telephone circuit modifications. However, by 1947, with the appearance of the watershed papers of French and Steinberg (1947) and Beranek (1947), it was evident that articulation theory had the potential for much greater significance and broader application than was implied by these early goals.

The term articulation index (AI) was coined at the Bell Telephone Laboratories and first appeared in company memos as early as 1926 (French, 1926). It replaced the term quality index that Harvey Fletcher (1921) had proposed earlier. The new term better reflected the relationship between the index and what were then called articulation tests. Articulation tests were speech tests in which listeners were asked to identify speech sounds spoken by a caller under conditions of interest to the experimenter. The exact makeup of the articulation tests varied, but they usually consisted of a carrier sentence with a nonsense syllable test item at the end. Several callers would in turn utter the sentences via the test circuit to crews of six to eight listeners. The listeners would record what they heard phonetically. The circuit articulation was equal to the average proportion of the sounds (or syllables) heard correctly by the listeners. The articulation index was devised as an alternative to this procedure.

The articulation index is an index that describes the proportion of the total importance-weighted speech signal that is audible to the listener under specified conditions. Normally, the index’s value (ranging from 0 to 1) is derived from physical measurements, including principally the speech signal’s intensity level, the level of any noise that may be present, and the characteristics of the transmission system that delivers the speech and noise to the listener’s ear. Reverberation effects are sometimes included. Index values are often modified based on certain well-known performance characteristics of the human auditory system. Most commonly these would include pure-tone thresholds and cross-frequency band spread of masking. The negative effects of listening at high signal levels might also be included, as well as positive factors, such as the effect of the listener being able to see the talker’s face. The articulation index value can be used directly as an indicator of the relative quality of a communications system, or it can be used to predict average speech recognition success for the particular types of speech or speech elements under specified listening conditions through the use of an appropriate transfer function.

The early history of development of the articulation index is difficult to follow with certainty because much of this work was done at the Bell Telephone Laboratories and was described only in internal company memos and reports. Most was not published in the scientific literature for 10-25 years after it was carried out. Some was never published. Fortunately, copies of many of the original documents from this early period recently became available on a compact disk through the efforts of C. M. Rankovic and J. B. Allen (Rankovic and Allen, 2000).

The first credible effort to evaluate the effectiveness of a communication system based on physical measurements was that of H. Fletcher (1921). In an unpublished Western Electric Laboratory report, Fletcher pointed out that a suitable measure (index) of circuit quality must have the property of additivity. By this he meant that if a particular frequency range of speech heard alone has a quality value of \( Q_1 \) and if a second frequency range has a value of \( Q_2 \), then the value of both ranges heard at the same time should equal \( Q_1 + Q_2 \). It was clear that articulation test scores did not possess this
property (i.e., $A_1 + A_2 \neq A_{12}$). Therefore, Fletcher proposed an intermediate variable, related to articulation, that would at least approximate this additivity property. He also proposed methods to derive the index’s value for a circuit from measures of received speech intensity, frequency distortion, room and line noise, “asymmetric distortion,” and other factors (Fletcher, 1921).

The term articulation index first appeared in the published literature in the landmark paper by Bell Telephone Laboratory scientists N. R. French and J. C. Steinberg in 1947. However, the term appears in internal Bell Labs documents as a replacement for quality index as early as 1926. The expression quality theory, however, continued in use for several more years before finally being replaced by articulation theory.

In a parallel development in England, John Collard, working for International Telephone and Telegraph, published a detailed description of an index with properties similar to those of the articulation index in a series of papers beginning in 1930. He called his index band articulation (or frequently just “the new unit”). Speech test scores were called sound articulation. By 1939 Collard had created a mechanical band articulation calculator (Pocock, 1939) that included many of the features of more recent articulation index methods. It is unclear whether the Bell Telephone Laboratory scientists were aware of Collard’s work, but it is not cited in either their published or unpublished reports.

During World War II, Fletcher, then director of physical research at Bell Telephone and chairman of the National Defense Research Committee (NDRC), provided a description of at least some of the articulation index methods that had been developed at Bell Labs to Leo Beranek, then at Harvard University (Allen, 1996). Beranek was working on methods for improving communications for aircraft pilots as a part of the war effort under a contract from the NDRC. Allen (1996) reports that following the war, Beranek persuaded the Bell Labs group to finally publish a description of their work on the articulation index. The classic 1947 paper by N. R. French and J. C. Steinberg was the result. This paper was soon followed by Beranek’s frequently cited paper on the articulation index (Beranek, 1947). These two papers together were to play highly influential roles in the future of the articulation index.

In 1950 Fletcher and his long-time Bell Labs associate Rogers Galt published a detailed description of their conception of the articulation index. Their goal in this paper was to cover a broader range of conditions than had ever been attempted before. “Telephony” was defined by the authors as referring to “any talker-listener combination” (p. 90). The results were seen as applicable to sound recording and reproduction systems, public address systems, and even hearing aids. In 1952, Fletcher further extended the applications to persons with hearing loss. Unfortunately, Fletcher and Galt’s attempt to account for all aspects of the problem resulted in a conceptualization and method that was too complex for most people to understand, and it was seldom used. Recently, there has been a renewed interest in the method (Rankovic, 1997, 1998). An available computer program implementing the procedure by H. Müsch (2001) now makes its practical use feasible.

The next major steps in the history of the articulation index were taken by Karl Kryter. In two landmark papers (Kryter, 1962a, 1962b), he described and validated a comprehensive method for calculating the articulation index under a broad range of conditions. This work was based directly on the publications of Beranek (1947) and French and Steinberg (1947). Kryter’s methods became even more influential in 1969 when they were adopted, virtually intact, by the American National Standards Institute (ANSI) as the American National Standard Methods for the Calculation of the Articulation Index (ANSI S3.5-1969). For some 30 years this method quite literally defined the articulation index.

An important development taking place in Europe during this period was that of the speech transmission index by T. Houtgast and H. J. M. Steeneken (Houtgast and Steeneken, 1973; Steeneken and Houtgast, 1980). Similar in many ways to the articulation index of French and Steinberg (1947), the speech transmission index added a unique method for incorporating and combining the effects of noise and reverberation into the calculation through measurements of the modulation transfer function (MTF). At first called the weighted MTF, the speech transmission index has found relatively widespread use in the area of architectural acoustics in an abbreviated form known as the rapid speech transmission index, or RASTI.

In 1997 ANSI published the American National Standard Method for Calculation of the Speech Intelligibility Index (ANSI S3.5-1997). This renamed standard is the direct successor of articulation index standard ANSI S3.5-1969 and it is similar to the earlier standard in basic concept. However, there are many differences in detail. One of the more obvious is the change in the name of the index to speech intelligibility index. The new name finally severs the connection with the now obsolete term, articulation test, and also avoids confusion between its abbreviation, AI, and the newer but more general use of this abbreviation to mean artificial intelligence.

Of the procedural differences, one of the more fundamental ones concerns the frequency importance function. The frequency importance function describes the relative importance of different frequency regions along the frequency scale for speech intelligibility. ANSI S3.5-1997 provides a function for average speech, but it also provides different functions for particular types of speech. This change from the earlier standard implies that frequency importance is significantly dependent on the characteristics of the speech material and not only dependent on the characteristics of the auditory system. In addition, ANSI S3.5-1997 provides for calculations based on the auditory critical band, uses newer methods for calculating spread of masking, includes a speech level distortion factor, uses different speech spectra for raised, loud, and shouted speech, and provides for use of the modulation transfer function methods of Houtgast and Steeneken (1980). To ensure accuracy, a computer
program implementing the S3.5-1997 method was made available. This program lacks a user-friendly interface but provides an invaluable test of other implementations of the method.

—Gerald A. Studebaker

References


Further Readings


Speech Tracking

Speech Tracking (sometimes called Connected or Continuous Discourse Tracking) is a procedure developed by De Filippo and Scott “for training and evaluating the reception of ongoing speech” (De Filippo and Scott, 1978, p. 1186). In the Speech Tracking procedure, a talker (usually a therapist or experimenter) reads from a prepared text, phrase by phrase, for a predetermined time period, usually five or ten minutes. The task of the receiver (the person with hearing loss) is to repeat exactly what the talker said. If the receiver does not give a verbatim response, the talker applies a strategy to overcome the breakdown. This can take the form of simple repetition, the use of clue words, or a paraphrase of the original segment. The therapist goes on to the next phrase only when the receiver is able to repeat correctly every word of the original segment.

At the end of the time period, the number of words repeated correctly is counted and divided by the time elapsed to derive the receiver’s Tracking Rate, expressed in words per minute (wpm). For example, if a receiver was able to repeat 150 words in five minutes, his Tracking Rate would be 30 wpm. The Tracking Rate represents the time taken for the text to be presented and repeated. Estimates of Tracking Rate for people with normal hearing vary, but it is generally recognized that a rate of around 100 wpm (De Filippo, 1988) is obtained if the same presentation and response rules are followed.

Since its introduction, Speech Tracking has also been used extensively in investigations evaluating the effectiveness of sensory aids for people with hearing loss. These have included studies of cochlear implants (Robbins et al., 1985; Levitt et al., 1986), tactile aids (Brooks et al., 1986; Cowan et al., 1991; Plant, 1998), and direct contact tactile approaches such as Tadoma (Reed et al., 1992) and Tactiling (Plant and Spens, 1986). These studies usually compare a receiver’s Tracking Rate in two or more presentation conditions such as aided and unaided lip reading. For example, Plant (1998) looked at a subject’s Speech Tracking performance with materials presented via lip reading alone and lip reading supplemented by the Tactaid 7 vibrotactile aid. After about 30 hours of testing and training with Speech Tracking, the subject’s mean Tracking Rates in the two presentation conditions were 33.7 wpm for lip reading alone and 46.1 wpm for lip reading supplemented by the Tactaid 7.

Despite the widespread acceptance of Speech Tracking in research projects, its use as an evaluative tool has been severely criticized by Tye-Murray and Tyler (1988). These researchers cited a number of extraneous variables that they believed were extremely difficult to control. These included the characteristics of the speaker (segment selection, ability to use cues to overcome blockages, speaking style, articulatory patterns, etc.), the receiver (degree of assertiveness, language proficiency, motivation, etc.), and the text (degree of syntactic complexity, vocabulary, etc.). Hochberg, Rosen, and Ball (1989), for example, found that text complexity could greatly influence Tracking Rate. In a study using the same talker/receiver pairs, they found that Tracking Rates varied from 62.9 wpm for “easy” materials (controlled vocabulary readers designed for English-as-a-second-language learners) to 29.5 wpm for “difficult” materials (popular adult fiction).

Tye-Murray and Tyler (1998) believed that these factors made Speech Tracking unsuitable for across-subject test designs. They did, however, feel that with some modifications the procedure could be used for within-subject test designs. These recommendations included an insistence on a verbatim response, the use of only one speaker, training of the speaker/receiver pairs, the use of appropriate texts, and limiting the repair strategies used to repetition and writing down a blocked word after three repeats.

A number of groups (for example, Boothroyd, 1987; Pichora-Fuller and Benguerel, 1991; Dempsey et al., 1992) have attempted to make the technique more suitable for evaluative purposes through the use of computer-controlled recorded materials. These approaches ensure that the problems created by sender differences are controlled and minimized. Although promising, none of these systems have become widely accepted and used.

The KTH Speech Tracking Procedure (Gnosspeilius and Spens, 1992; Spens, 1992) represented another attempt to more closely control the approach. This computer-controlled modification used live-voice presentations, but the segment length was predetermined and only one repair strategy—repetition—was allowed. The written form of any word repeated three times was automatically presented to the receiver via an LED display. At the end of a Speech Tracking session the program automatically calculated a number of measures including Tracking Rate, Ceiling Rate (time taken when all words in a segment were correctly repeated after only one presentation), and the Proportion of Blocked Words (the total number of words that had to be repeated divided by the total number of words in the session). This approach has been used in a number of studies (for example, Plant, 1998; Ronnberg et al., 1998) but has not gained widespread acceptance.
Speech Tracking has also been widely used as a training technique for use with adults (Plant, 1996) and children (Tye-Murray, 1998) with profound hearing loss. When used for training, modifications can be made to provide receivers with practice in the use of repair strategies. Owens and Raggio (1987), for example, provided receivers with a list of directives they could use when they did not correctly repeat a segment. The receiver could ask the sender to:

1. Say that again.
2. Say it another way.
3. Spell that word.
4. Write that word.
5. Spell an important word.
6. Write an important word.

Lunato and Weisenberger (1994) looked at the comparative effectiveness of four repair strategies in Speech Tracking. The repair strategies used were:

1. Verbatim repetition of a word or phrase.
2. The use of antonyms or synonyms as clues to the identity of blocked words.
3. Providing the receiver with phoneme-by-phoneme correction of blocked words.
4. Providing context by moving forward or backward in the text.

These researchers reported that strategy 1 yielded the highest Tracking Rates and strategy 2 the lowest.

When Speech Tracking is being used for training, compromises can also be made in the receiver’s response patterns. Owens and Raggio (1987), for example, argued for the use of nonverbatim responses in training sessions using Speech Tracking. While acknowledging the importance of a verbatim response for test purposes, they felt that in training it may be better to provide the receiver with practice in picking up the gist of the message rather than expecting absolute identification at all times.

Although Speech Tracking has become a widely used training procedure, there are some people with hearing loss for whom it is unsuitable. These include people with very poor speech reception skills, resulting in Tracking Rates of less than 20 wpm. At these levels receivers find the task extremely difficult and stressful. Others for whom the technique may be unsuitable include people with poor speech production skills. In such cases the sender may be unable to determine reliably whether the receiver gave the correct response. This necessitates the use of a written or a signed response, which serves to greatly reduce the Tracking Rate.

Plant (1996, 1989) developed a modified version of Speech Tracking designed to be used with such cases. Simple stories are divided into parts, each consisting of 200 words. Each part is in turn divided into short segments ranging in length from 4 to 12 words. The segments are presented for identification, and the receiver is asked to repeat as many words as he or she can. The receiver can use repair strategies to obtain additional information if he or she experiences difficulties. The receiver is then scored for the number of words correctly identified and shown the written form of the segment. At the end of each part, the percentage of correct responses is calculated, based on the number of words correctly identified. This approach provides the receiver with immediate feedback on the correctness of her or his response and ensures that she or he is able to benefit from ongoing contextual information.

Speech Tracking is an innovative technique that can be used for both testing and training the speech perception skills of people with hearing loss. When used for testing, however, a precise protocol must be followed to minimize the effects of sender, receiver, and text variables. In training a less rigid approach can be used, and the approach may be modified to include practice in the use of repair strategies.

—Geoff Plant

References


Speechreading Training and Visual Tracking

Speechreading (lipreading, visual speech perception), a form of information processing, is defined by Boothroyd (1988) as a “process of perceiving spoken language using vision as the sole source of sensory evidence” (p. 77). Speechreading, a natural process in everyday communication, is especially helpful when communicating in noisy and reverberant conditions because facial motion in speech production may augment or replace degraded auditory information (Erber, 1969). Also, visual cues have been shown to influence speech perception in infants with normal hearing (Kuhl and Meltzoff, 1982), and speech perception phenomena, such as the McGurk effect (MacDonald and McGurk, 1978), demonstrate the influence of vision on auditory speech perception.

To understand language, the speechreader directs attention to, extracts, and uses linguistically relevant information from a talker’s face movements, facial expressions, and body gestures. This information, which may vary within and across talkers (for a review, see Kricos, 1996), is integrated with other available sensory cues, such as auditory cues, as well as knowledge about speech production and language in general to make sense of the visual information. However, the visual information may be ambiguous because many sounds look alike on the lips, are hidden in the mouth, or are co-articulated during speech production. In addition, expectations about linguistic context may influence understanding. Nevertheless some individuals are expert speechreaders, scoring more than 80% correct on words in unrelated sentences, and demonstrate enhanced visual phonetic perception (Bernstein, Demorest, and Tucker, 2000). Attempts to relate speechreading proficiency to other sensory, perceptual, and cognitive function, including neurophysiological responsiveness, have met with limited success (for a review, see Summerfield, 1992).

From a historical perspective, speechreading was initially developed in Europe as a method to teach speech production to young children with hearing loss. Until the 1890s it was limited to children and was characterized by a vision-only (unisensory) approach. Speechreading training was based on analytic methods, which encouraged perceivers to analyze mouth position to recognize sounds, words, and sentences, or synthetic methods, which encouraged perceivers to grasp a speaker’s whole meaning (Gestalt). O’Neill and Oyer (1961) reviewed several early distinctive methods that we adopted in America: Bruhn’s method (characterized by syllable drill and close observation of lip movements), Nitchie’s method (which shifted from an analytical to a synthetic method), Kinzie’s method (Bruhn’s classification of sounds plus Nitchie’s basic psychological ideas), the Jena method (kinesthetic and visual cues), and film techniques (Mason’s visual hearing, Markovin and Moore’s contextual systemic approach). Gagne´ (1994) reviewed present-day approaches: multimodal speech perception (integration of available auditory cues with those from vision and other modalities), computer-based activities (interactive learning using full-motion video), and conversational contexts (question-answer approach, effective communication strategies, and training for talkers with normal hearing to improve communication behavior).

When indicated, speechreading training is included in comprehensive programs of aural rehabilitation. At best, the post-treatment gains from speechreading training are modest, in the range of approximately 15%. Unfortunately, data on improvements related to visual speech perception training are limited, and little is known about the efficacy of various approaches. However, some individuals demonstrate significant gains. Walden et al. (1977) have reported an increase in the number of visually distinctive groups of phonemes (visemes) with which consonants are recognized following practice. The identification responses following practice suggest that the distinctiveness of visual phonetic cues related to place-of-articulation information is increased. Although it is not clear what factors account for such improvements in performance, these results provide evidence for instances of learning in which perception is modified. Other studies show that the results are variable, both in performance on speechreading tasks and in gains related to learning programs. Improvements observed by Masaro, Cohen, and Gesi (1993) suggest that repeated testing experience may be as beneficial as structured
training. Initial changes may be due to nonsensory factors such as increased familiarity with the task or improved viewing strategies. In contrast, Bernstein et al. (1991) suggest that speechreaders learn the visual phonetic characteristics of specific talkers after long periods of practice. Treatment efficacy studies may be enhanced by enrolling a larger number of participants, specifying training methods, using separate materials for training versus testing, evaluating asymptotic performance and long-term effects, determining whether the effects of the intervention are generalized to nontherapy situations, and designing studies to control for factors such as the motivation or test-taking behaviors of participants, as well as personal attention directed toward participants (for a review, see Gagné, 1994; Walden and Grant, 1993).

Current research has also focused on visual speech perception performance in psychophysical experiments. One research theme has centered on determining what regions of the face contain critical motion that is used in visual speech perception. Subjective comments from expert lipreaders suggest that movement in the cheek areas may aid lipreading. Data from Lansing and McConkie (1994) illustrate that eye gaze may shift from the mouth to the cheeks, chin, or jaw during lipreading. Results from Greenberg and Bode (1968) support the usefulness of the entire face for consonant recognition. In contrast, results from Ijsselstijl (1992) and Marassa and Lansing (1995) indicate that information from the lips and mouth region alone is sufficient for word recognition. Massaro (1998) reports that some individuals can discriminate among a small set of test syllables without directly gazing at the mouth of the talker, and Preminger et al. (1998) demonstrate that 70% of visemes in /a/ and /aw/ vowel contexts can be recognized when the mouth is masked. These diverse research findings underscore the presence of useful observable visual cues for spoken language at the mouth and in other face regions.

Eye-monitoring technology may be useful in understanding the role of visual processes in speechreading (Lansing and McConkie, 1994). It provides information about on-line processing of visual information and the tendency for perceivers to direct their eyes to the regions of interest on a talker’s face. By moving the eyes, a perceiver may take advantage of the densely packed, highly specialized cone receptor cells in the fovea to inspect visual detail (Hallet, 1986). Eye monitoring has been used to study a variety of cognitive tasks, such as reading, picture perception, and face recognition, and to study human-computer interaction (for a review, see Rayner, 1984). The basic data obtained from eye monitoring reveals sequences of periods associated with perception in which the eye is relatively stable (fixations) and high-velocity jumps in eye position (saccades) during which perception is inhibited. Distributions of saccadic information (“where” decisions) are quantified in terms of length and direction, and distributions of fixations (“when” decisions) are quantified in terms of duration and location. Experiments are designed to evaluate the variance associated with cognitive processes. For exam-
ple, distributions of fixation duration and of saccade length differ across cognitive tasks such as reading versus picture perception. Various types of instrumentation are available for eye-monitoring research, some of which include direct physical contact with the eyes to camera-based video systems that determine eye rotation characteristics based on changes in the location of landmarks such as the center of the pupil or corneal reflections, free of error induced by translation related to head movement (for a review, see Young and Sheena, 1975). Factors such as cost, accuracy, ease of calibration, response mode, and demands of the participants and experimental task must be considered in selecting an appropriate eye-monitoring system. An example of a system used in speechreading research is shown in Figure 1. The system is used to record the eye movements of the perceiver and to obtain a detailed record of the sequence and duration of fixations. A scan plot and sample record of eye movements are shown in Figure 2. Simultaneously, measurements are made of the accuracy of perception, efficiency of processing, or judgment of stimulus difficulty. For interpretation, the eye movement records are linked to the spatial and temporal characteristics of face motion for each video frame or speech event (e.g., lips opening).

Results from eye-monitoring studies demonstrate that speechreaders make successive eye gazes (fixations) to inspect the talker’s face or to track facial motion. The talker’s eyes attract attention prior to and following speech production (Lansing and McConkie, 2003), and in the presence of auditory cues (Vatikiotis-Bateson, Eiigisti, and Munhall, 1998). If auditory cues are not available, perceivers with at least average proficiency attend to the talker’s mouth region for accurate sentence perception (Lansing and McConkie, 2003). However, some gazes are observed toward the regions adjacent to the lips as well as toward the eyes of the talker. Similarly for word understanding, speechreaders direct eye gaze most often and for longer periods of time toward the talker’s mouth than toward any other region of the face. Motion in regions other than the mouth may increase signal redundancy from that at the mouth and afford a
natural context for observing detailed mouth motion. Task characteristics also influence where people look for information on the face of the talker (Lansing and McConkie, 1999). Eye gaze is directed toward secondary facial cues, located in the upper part of the face, with greater frequency for the recognition of intonation information than for phonemic or stress recognition. Phonemic and word stress information can be recognized from cues located in the middle and lower parts of the face.

Finally, new findings from brain imaging studies may provide valuable insights into the neural underpinnings of basic processes in the visual perception of spoken language (Calvert et al., 1997) and individual differences in speechreading proficiency (Ludman et al., 2000). Although preliminary results have not yet identified speechreading-specific regions, measures in perceivers with normal hearing indicate bilateral activation of the auditory cortex for silent speechreading (Calvert et al., 1997). Results from measures in perceivers with congenital onset of profound bilateral deafness (who rely on speechreading for understanding spoken language) do not indicate strong left temporal activation (MacSweeney et al., 2002). Functional magnetic resonance imaging may prove to be a useful tool to test hypotheses about task differences and the activation of primary sensory processing areas, the role of auditory experience and plasticity, and neural mechanisms and sites of cross-modal integration in the understanding of spoken language.

Continued study and research in the basic processes of speechreading are needed to determine research-based approaches to intervention, the relative advantages of different approaches, and how specific approaches relate to individual needs. Additional insight into the basic processes of visual speech perception is needed to develop and test a model of spoken word recognition that incorporates visual information, to optimize sensory-prosthetic aids, and to enhance the design of human-computer interfaces.

—Charissa R. Lansing

References


Further Readings


Suprathreshold Speech Recognition

Suprathreshold refers to speech presented above the auditory threshold of the listener. Speech recognition is generally defined as the percentage of words or sentences that can be accurately heard by the listener. For example, a patient who could correctly repeat 40 out of 50 words presented would have 80% speech recognition. Because speech is a complex and continually varying signal requiring multiple auditory discrimination skills, it is not possible to accurately predict an individual’s speech recognition from the pure-tone audiogram (Marshall and Bacon, 1981). Measurement of suprathreshold speech recognition allows clinicians to assess a patient’s speech communication ability in a controlled and systematic manner. The results can help clinicians distinguish between different causes of hearing loss and plan and evaluate audiological rehabilitation programs.

Speech is a complex acoustic signal that varies from moment to moment: from shouting to whispering, from clear speech in quiet to difficult to understand speech in high ambient noise. Figure 1 shows the expected frequency and intensity of speech sounds for speech spoken at a conversational level in quiet. In general, the vowel sounds contain lower frequency information and are more intense, while consonant sounds contain higher frequency information and are produced at a lower intensity level. Which sounds are audible depends on the listener’s audiometric thresholds as well as on the speaker and the level at which the words are spoken. For example, the sounds of shouted speech would be shifted to a higher intensity level and have a slightly different pattern across frequency (Olsen, 1998). Figure 1 also shows audiograms for two different listeners. The first listener has a moderately severe hearing loss. This listener would likely hear few, if any, conversational speech sounds and would be expected to have very poor speech recognition. The second listener has normal hearing in the low frequencies falling to a mild hearing loss in the high frequencies. This listener might hear some, but not all, of the speech sounds. The inability to hear high-frequency consonants would likely result in less than 100% speech recognition for a conversational-level signal.

Most commonly, the material used to measure speech recognition is a list of monosyllabic words. Typically each word is preceded by a carrier phrase, such as “Say the word——.” Most available monosyllabic word lists are open set; that is, the listener is not restricted to a predetermined list of possible responses. A number of standard lists have been developed with vocabulary levels appropriate for adults (Hirsh et al., 1952; Tillman and Carhart, 1966) or children. The lists exclude words that would be unfamiliar to most people, and word selection is balanced to maintain similar levels of difficulty across lists. Additionally, each list is phonetically balanced; that is, the sounds in the words occur in the same proportion as in everyday speech. Test sensitivity is enhanced by using a larger number of items, such as 50 instead of 25 words (Thornton and Raflin, 1978).

Sentence tests are also available. These tests are more like “real speech” and thus presumably able to provide a closer estimate of real-life communication performance. However, sentence tests incorporate additional factors besides simple audibility. With sentences, the listener may be relying on linguistic, prosodic, or contextual cues in addition to auditory information. To limit contextual cues, sentence lists are available that use neutral context (Kalikow, Stevens, and Elliot, 1977) or linguistically meaningless word combinations (Speaks and Jerger, 1965). Sentences also place greater demands on higher-level processes such as auditory memory, which may be a particular problem in older listeners (Chmiel and Jerger, 1996).

For standard clinical testing, the participant is seated in a sound booth and listens to speech presented to one ear at a time through earphones. Speech may also be presented through a speaker, although this does not provide information specific to each ear. Such sound field testing can be used to quantify the effects of amplification. Recorded speech materials are preferred for consistency, although speech recognition tests are also administered using monitored live voice, during which the tester speaks to the participant over a microphone while monitoring her vocal strength. The entire list of words is presented at the same level. After each word, the participant responds by repeating or writing the word. The speech recognition score is then expressed as the percentage of correct words at the given presentation level in each ear.

Although most often presented in quiet, these materials may also be administered in noise. Many recordings include a background of multitalker babble that mimics a more realistic listening situation; this increases the degree to which test results characterize the listener’s...
everyday communication abilities. For example, listeners with sensorineural loss require a more favorable signal-to-noise ratio than listeners with normal hearing or conductive hearing loss (Dubno, Dirks, and Morgan, 1984).

When administering and interpreting suprathreshold speech recognition tests it is important to consider not only the test environment but also the physical, linguistic, cognitive, and intellectual abilities of the listener. If the listener is unable to respond verbally or in writing, tests are available where the listener can choose among a set of picture responses (Ross and Lerman, 1970). Although most often used with children, these tests are also appropriate for adults with spoken word deficits, including dysarthria or apraxia. One limitation of such closed-set tests is that the chance of guessing correctly is higher when only a fixed number of choices is available. However, scoring accuracy may be higher than with open-set tests because there are fewer chances for misinterpretation of the response. For listeners who are not proficient in English, recorded materials are available in a number of other languages (provided, of course, that the tester has sufficient knowledge of the test language to interpret responses).

An important consideration is the presentation level. If multiple levels are tested, the percentage correct increases with increasing presentation level in a characteristic pattern (Fig. 2). This is referred to as the performance intensity (PI) function. The rate of improvement depends on the test material as well as patient characteristics. Easier material (e.g., sentences containing contextual cues) results in a greater rate of improvement with increases in level than more difficult material (e.g., nonsense words). The presentation level at which the listener achieves a highest score is referred to as the PB max, or maximum score for phonetically balanced words. A normal-hearing listener typically achieves 100% speech recognition at levels 30–40 dB above the SRT. Sensorineural hearing loss may restrict the PB max to below 100%. Listeners with conductive hearing loss generally achieve 100% recognition, although they require a higher presentation level than would a normal-hearing listener. PI functions for listeners with retrocochlear loss may demonstrate disproportionately low scores as well as a phenomenon called rollover, in which performance first improves with increasing presentation level, and then degrades as the presentation level continues to increase.

In the clinic, speech recognition testing is often done at only one or two levels in each ear to minimize test time. One common approach is to select one or more levels relative to the speech reception threshold. Selection of the specific presentation level is generally based on providing adequate speech audibility, particularly at frequencies containing important consonant information. An alternative approach is to present speech at the level the listener deems most comfortable. Because the listener’s most comfortable level may not be the same level at which she obtains a maximum score, testing exclusively at the most comfortable level can lead to erroneous conclusions about auditory function (Ulrich and Grimm, 1976; Beattie and Warren, 1982).

In summary, measurement of suprathreshold speech recognition is an important part of an audiometric examination. Test results can be affected by a number of factors, including the participant’s pure-tone sensitivity, the amount of distortion produced by the hearing loss, the presentation level of the speech, the type of speech material, the presence or absence of background noise, and even the participant’s age. A detailed understanding of these factors is important when interpreting test results and drawing conclusions about an individual’s overall communication ability.

—Pamela E. Souza

References


**Further Readings**


**Temporal Integration**

The term *temporal integration* (TI) refers to summation of stimulus intensity during the duration of the stimulus. As duration increases, a sensation like loudness increases, or the sound level at which the stimulus can be detected decreases. The stimuli may be various types of signals, such as tones or bands of noise. Similarly, short succeeding stimuli can combine their energies and provide a lower detection level than individual stimuli. The

**Figure 1.** Temporal integration curves according to the functions shown in the legend. In curves $A_1$ and $A_2$, the time constant $\tau = 300$ ms; in $A_1$, exponent $m = 1$; in $A_2$, $m = 0.8$. In $B_1$, $\tau = 300$ ms; in $B_2$, $\tau = 100$ ms. The value of $\tau$ = 300 ms is indicated by a mark on the abscissa.

TI has a time limit. For a stimulus longer than this limit, the loudness, or the detection (threshold) level, remains relatively constant.

Interest in studying TI is fueled by the need to understand auditory processing of speech—a signal that, by its nature, changes rapidly in time. Better understanding of the temporal characteristics of hearing should help us improve means for enhancement of speech communication in unfavorable listening environments, and of listeners with impaired hearing.

Graphs of the relationship between the stimulus duration (plotted on the horizontal coordinate, usually in milliseconds with a logarithmic scale) and the intensity level at the threshold of hearing (plotted on the vertical coordinate in decibels, dB) are called temporal integration curves (TICs). Examples of TICs are shown in Figure 1. The detection intensity level first declines as the stimulus duration increases and then, beyond a time limit called the critical duration, remains constant. The magnitude of TI can be expressed by the difference between the detection levels of long and short signals. The rate of decline of TICs is represented by slopes of the curves, which too are often used as indicators of TI magnitude. These slopes (they are negative) are usually expressed as the ratio of the change of level (in dB) per tenfold increase in signal duration [(L2–L1)/dec], or per doubling of signal duration. The slopes of the TICs and the values of the critical duration represent summary characteristics of TI. (The critical duration depends on a time constant $\tau$, a parameter in formulas describing TICs.)

**Factors Affecting TI**

The slope of the curves and the time constant depend on various factors, such as signal frequency, status of hear-
Temporal Integration

The effect of signal frequency on TI is pronounced (Gerken, Bhat, and Hutchison-Clutter, 1990) and depends on signal duration. TI, as well as τ, is greater at lower frequencies than at higher ones (e.g., Fastl, 1976; Nábelek, 1978; Florentine, Fastl, and Buus, 1988). At low frequencies and signal durations below 10 ms, the TIC slopes were found to be up to −15 dB/dec (e.g., Green, Birdsall, and Tanner, 1957). At frequencies between 1 and 8 kHz and at signal durations between 20 and 100 ms, the slopes are between −10 and −8 dB/dec (e.g., Zwislocki, 1969; Gerken, Bhat, and Hutchison-Clutter, 1990). The steeper slopes at short signal durations as compared to slopes at longer signal durations are attributed to the loss of contribution of some energy due to spectral broadening, or “splitter.” When the frequency during the signal is not constant but is increasing, the slope between 20 ms and 80 ms of signal duration is smaller, about −9 dB/dec, than when the frequency is decreasing—about −13 dB/dec (Nábelek, 1978). For broadband masking conditions, the values of TI for constant tones are similar to those without masking, but some influence of the level of the masker was observed. This influence depends on signal duration. For signal durations between 2 and 10 ms, the TICs at medium masker levels are steeper than at low or high masker levels, and for signal durations over 20 ms, the TI values are not affected by the masker level (Oxenham, Moore, and Vickers, 1997). Formby et al. (1994) investigated the influence of bandwidth of a noise signal masked by an uncorrelated broadband noise on TI and τ. They found that both TI and τ were related inversely to the bandwidth, if the bandwidth was greater than the critical band of hearing (CB), and were relatively invariant if the bandwidth was smaller than the critical band. For gated signal and masker, Formby et al. (1994) identified at least three cues for signal detection: (1) a relative timing cue, (2) a spectral shape cue, and (3) a traditional energy cue. The timing and spectral shape cues count most for the shortest (10 ms) and narrowest (bandwidth = 63 Hz) signals, respectively. When the signal is a series of tone pulses, and not single bursts, the change of time interval between the pulses produces smaller change of TI than the change in duration of single bursts (Carlyon, Buus, and Florentine, 1990).

Listeners with hearing impairment generally show less temporal integration than listeners with normal hearing (e.g., Watson and Gengel, 1969; Gerken, Bhat, and Hutchison-Clutter, 1990). No effect of level of a broadband masker was found for listeners with impaired hearing at any signal duration.

 Loudness increases when the duration of a short signal increases. When the signal level changes, the TI for loudness changes; however, the change is not monotonic (Buus, Florentine, and Poulsen, 1999). The change is greatest at moderate sensation levels and depends on signal duration. The effect of signal level on TI of loudness is greater at short than at long signal durations.

Donaldson, Viemeister, and Nelson (1997) found that TICs for electrical stimulation with the Nucleus-22 electrode cochlear implant were considerably less steep than −8 dB/dec typically observed with acoustical stimulation. The slopes varied widely across subjects and across stimulated electrodes. When Shannon and Otto (1990) used a device called the auditory brainstem implant (ABI) and positioned its electrodes near the cochlear nucleus of listeners, they obtained only a shallow TIC over the range of 2- to 1000-ms signal duration.

Models

A number of models for temporal integration have been proposed. The theoretical foundations for the mathematical description of TI are either deterministic or probabilistic. Deterministic models include power function models or exponential function models. One of the deterministic models is described mathematically by the power function \( t(I, I_\tau) = I_\tau \tau = \text{const} \) (Hughes, 1946), or in its more general form by \( I^m = C \) (Green et al., 1957). In these equations \( t \) is the stimulus duration, \( I_\tau \) is the threshold intensity at \( t \), \( I_\tau \) is the threshold intensity for very long stimuli, \( \tau \) is the time constant of the integration process, \( m \) is the power function exponent, and \( C \) is a constant. The exponent \( m \) determines the slope of the curves (A1 and A2 in Fig. 1). The slope \( -3 \text{ dB/doubling} \) or \( -10 \text{ dB/dec} \) corresponds to \( m = 1 \). Another model is the exponential one \( I/I_\tau = 1/(1 - e^{-t/\tau}) \) (Feldkeller and Oetinger, 1956; Plomp and Bouman, 1959). The curves B1 and B2 in Figure 1 correspond to this equation.

Zwislocki (1960) developed a temporal summation theory for two pulses separated in time and proposed a theory of TI for loudness (Zwislocki, 1969). In his model it is assumed that (1) a linear temporal integrator (with a time constant on the order of 200 ms) exists in the central nervous system; (2) a nonlinear transformation that produces compression precedes the temporal summation; and (3) neural excitation decreases exponentially with a short time constant at the input to the integrator that summates the central neural activity. (This last assumption indicates that the term temporal integration should not be interpreted as the integration of acoustic energy per se.) Attempts to resolve an apparent discrepancy between high temporal resolution of hearing and long time constants of temporal integration have led to a number of models employing short integration times (e.g., Penner, 1978; Oxenham, Moore, and Vickers, 1997). Viemeister and Wakefield (1991) have not considered this discrepancy to be a real problem. Their model is based on a statistical probability approach and assumes multiple sampling. Taking their own data into account, Viemeister and Wakefield conclude that power integration occurs only for pulses separated in time by less than about 5 ms, and that therefore their data are inconsistent with the classical view of TI involving long-term integration. However, they find the data to be compatible with the notion that the input is sampled at a fairly high rate and the obtained samples (or “looks”) are stored in memory; while in the memory, the “looks”
can be selectively accessed, weighted, and otherwise processed.

Dau, Kollmeier, and Kohlrausch (1997) proposed a multichannel model. They describe the effects of spectral and temporal integration in amplitude-modulation detection for a stochastic noise carrier. The model is based on the concept of a modulation filter-bank. To integrate information across frequency, the detection process in the model combines cues from all filters with an optimal decision statistic. To integrate information across time, a “multiple-look” strategy, similar to that proposed by Viemeister and Wakefield (1991), is realized within the detection stage of the model. The temporal integration involves a template that provides the basis for the optimal detector of the model. The length and the time constant of the template are variable: they change according to the task which the listener has to perform.

Although an extensive knowledge of temporal integration has been attained, many aspects of TI await further investigation. For example, evidence of some TI mechanism at a higher stage of the auditory pathway was found by Uppenkamp, Fobel, and Patterson (2001) when they compared the perception of short-frequency sweeps and the physiological response to them in the brainstem. The improved understanding of TI should provide a sounder basis for the development of means for securing better speech communication in general and for listeners with special problems, like those with cochlear implants, in particular. Presently, TI studies are not limited to traditional topics but also cover higher levels of the brain, like the role of TI in establishing neural representations of phonemes (Tallal et al., 1998), and investigation of an association between a deficient TI and mental disturbances in schizophrenia (Haig et al., 2001; Michie, 2001).

Acknowledgments

Many thanks to Assoc. Prof. Lana S. Dixon for her help in securing pertinent references.

See also CLINICAL DECISION ANALYSIS; COCHLEAR IMPLANTS; MASKING; TEMPORAL RESOLUTION.

—Igor V. Nábelek

References


Temporal Resolution

Sensory systems function as change detectors in many respects. They quickly adapt to steady-state stimuli and are easily excited by the introduction of novel stimuli. The pattern of changes in an acoustic stimulus conveys information about the nature of the sound source and the message being transmitted by the sender. Therefore the identification, discrimination, and interpretation of acoustic events depend on the ability of the auditory system to faithfully encode the temporal features of those events. This ability to respond to changes in an acoustic stimulus has been termed temporal resolution. Although most natural acoustic signals are characterized by changes in intensity as well as changes in the acoustic spectrum over time, investigations of temporal resolution have focused on intensity variations in an attempt to separate purely temporal from spectro-temporal resolving capabilities (see Auditory Scene Analysis). Temporal resolution is limited by auditory inertia resulting from mechanical and/or electrophysiological transduction processes. Such a limitation effectively smoothes or attenuates the intensive changes of a stimulus, which reduces the salience of those changes. Impaired temporal resolution may be conceptualized as an increase in this smoothing process, and thus a loss of temporal information.

The influence of hearing impairment on temporal resolution depends on the site of lesion. For example, conductive hearing loss is often modeled as a simple attenuation characteristic and thus should not alter temporal resolution, given sufficient stimulus levels. Damage at the level of the cochlea, however, involves more than attenuation. Reduced outer hair cell function is associated with a reduction in sensitivity, frequency selectivity, and compression at the level of the basilar membrane. Each of these might influence the perception of intensity changes. For example, a loss of basilar membrane compression might provide a more salient representation of intensity changes and thus lead to improved performance on temporal resolution tasks involving such changes. Reduced frequency selectivity is analogous to broadening of a filter characteristic, which is associated with a shorter temporal response. This too might lead to improved temporal resolution. A loss of inner hair cell function, however, would reduce the quality and amount of information transmitted to the central auditory pathway, and might therefore lead to poor coding of temporal features. The altered neural function associated with a retrocochlear lesion may also lead to a less faithful representation of the temporal features of a sound.

Numerous techniques have been used to probe temporal resolution abilities; however, the two most common techniques are temporal gap detection and amplitude modulation detection (Fig. 1). Following the

![Figure 1](image-url)
neural hearing loss (Irwin, Hinchcliff, and Kemp, 1981). At low stimulus levels. At higher stimulus levels, gap thresholds for broadband noise are longer than normal.

normal gap detection thresholds. For patients with conductive or sensorineural hearing loss, gap detection thresholds for broadband noise are longer than normal at low stimulus levels. At higher stimulus levels, gap thresholds are within normal limits for conductive loss but remain longer than normal for listeners with sensorineural hearing loss (Irwin, Hinchcliff, and Kemp, 1981).

To gauge temporal resolution in different frequency regions, one may measure gap detection thresholds using band-limited noise. Results from listeners with normal hearing reveal that gap thresholds improve with increasing stimulus level up to about 30 dB sensation level (e.g., Buus and Florentine, 1983) and improve with increasing noise bandwidth (e.g., Shailer and Moore, 1983; Eddins, Hall, and Grose, 1992), but vary little with frequency region when noise bandwidth (in Hz) is held constant (e.g., Eddins et al., 1992). With hearing loss of cochlear origin, gap detection is often worse than normal using band-limited noise (e.g., Fitzgibbons and Wightman, 1982; Fitzgibbons and Gordon-Salant, 1987); however, this is not true for all listeners with cochlear hearing loss (e.g., Florentine and Buus, 1984; Glasberg and Moore, 1989; Grose, Eddins, and Hall, 1989). Thus, cochlear hearing loss does not necessarily result in poorer than normal temporal resolution.

Temporal gap detection thresholds measured for sinusoidal stimuli do not vary substantially with stimulus frequency from 400 to 2000 Hz, but increase substantially at and below 200 Hz (e.g., Shailer and Moore, 1987; Moore, Peters, and Glasberg, 1992). Although listeners with hearing impairment may have worse than normal gap detection thresholds for noise stimuli, gap detection thresholds for tonal stimuli are normal when compared at equivalent sound pressure levels and are better than normal at equal sensation levels (Moore and Glasberg, 1988). One theory consistent with these results is that gap detection is limited to some extent by the inherent fluctuations in narrow-band noise, and this effect may be accentuated by the loudness recruitment of some hearing-impaired listeners (e.g., Moore and Glasberg, 1988). Sinusoids, having a smooth temporal envelope, would not be subject to such a limitation. This leads to the possibility that temporal resolution per se may not be adversely affected by cochlear hearing loss (e.g., Moore and Glasberg, 1988). If gap detection is influenced by loudness recruitment, then one would expect a relationship between gap detection and intensity resolution. Indeed, gap detection for sinusoids is correlated with intensity resolution for sinusoids (Glasberg and Moore, 1989) and gap detection for band-limited noise is correlated with intensity resolution for band-limited noise (Eddins and Manegold, 2001). This highlights the potential role of intensity resolution in a task such as gap detection. Poor gap detection thresholds may result from poor intensity resolution, poor temporal resolution, or a combination of the two. Listeners with cochlear implants offer a unique perspective on temporal resolution in that the auditory periphery, save for the auditory nerve, is bypassed. Gap detection in such listeners, using electrical stimulation via the implant, is as good as gap detection for listeners with normal hearing using acoustic stimulation (e.g., Shannon, 1989; Moore and Glasberg, 1988). This is consistent with the notion that gap detection may not be strongly dependent upon cochlear processes.

While gap detection thresholds may be strongly influenced by a listener’s intensity resolution, the amplitude modulation detection paradigm provides an opportunity to separate the effects of intensity resolution from temporal resolution. A modulation detection threshold is obtained by determining the minimum depth of modulation necessary to discriminate an unmodulated from a sinusoidally amplitude-modulated stimulus. With this technique, temporal resolution can be more completely described as the change in modulation threshold over a range of fluctuation rates (modulation frequencies). With the assumption that intensity resolution does not vary with modulation frequency, a separate index of intensity resolution may be obtained from modulation detection thresholds at very low modulation frequencies. If loudness recruitment associated with cochlear hearing loss has a negative influence on gap detection in narrow-band noise, as suggested above, then one might predict that recruitment would enhance the perception of fluctuations introduced by amplitude modulation. Using broadband noise carriers, this does not seem to be the case. Modulation detection thresholds for listeners with cochlear hearing loss may be normal or worse than normal, but are not better than normal (Bacon and Vie-meister, 1985; Bacon and Gleitman, 1992). Similarly, modulation detection using band-limited noise is not worse than normal in listeners with cochlear hearing loss (e.g., Moore, Shailer, and Schooneveldt, 1992; Hall et al., 1998). Modulation detection with tonal carriers, however, tends to be better than normal in listeners with cochlear hearing loss, and the perceived depth of modulation appears to be related to the steepness of loudness growth (Moore, Wojtczak, and Vickers, 1996; Moore and Glasberg, 2001). This is quite different from amplitude-modulated noise stimuli, for which threshold does not seem to be related to loudness growth (Hall et al., 1998). As in the gap detection paradigm, there are marked differences between the results obtained with noise and tonal stimuli. Thus, it is possible that the relation between loudness growth and intensive changes is different for sinusoidal and noise stimuli.

Some listeners with cochlear pathology have worse than normal modulation detection using noise carriers,
as do listeners with Ménière’s disease (Formby, 1987), eighth nerve tumors (Formby, 1986), and auditory neuropathy (Zeng et al., 1999). Interestingly, listeners with cochlear implants perform as well as normal-hearing subjects on amplitude-modulation tasks (Shannon, 1992).

In summary, listeners with abnormal cochlear function often exhibit reduced performance on gap and modulation detection tasks with noise but not sinusoidal stimuli. Studies of temporal resolution using other experimental techniques have yielded results that are generally consistent with those discussed here. These results are consistent with an interpretation that cochlear pathology may not lead to reduced temporal resolution per se, but may lead to difficulty perceiving stimuli with pronounced, random intensity fluctuations (Grose et al., 1989; Hall and Grose, 1997; Hall et al., 1998). Although many hearing-impaired listeners perform as well as normal-hearing listeners on tasks involving temporal resolution, especially when stimuli are presented at optimal levels and have relatively smooth temporal envelopes, such listeners are likely to have difficulty in natural listening environments with fluctuating backgrounds. Thus, measures of gap and amplitude modulation detection using noise stimuli might have promise as predictors of communication difficulty in realistic environments.

—David A. Eddins

References


Further Readings


Tinnitus

*Tinnitus* is an auditory perceptual phenomenon that is defined as the conscious perception of internal noises without any outer auditory stimulation. Tinnitus may occur as a concomitant of practically all the dysfunctions that involve the human auditory system. Hence, damage to the middle ear, the cochlea, cranial nerve VIII (audiovestibular), and pathways in the brain from cochlear nucleus to primary auditory cortex all are likely candidates for explaining why tinnitus appears (Levine, 2001). A common distinction is between so-called objective tinnitus (somatosounds) and subjective tinnitus. In clinical settings, objective tinnitus represents a minority of cases. Examples of conditions related to objective tinnitus are spontaneous otoacoustic emissions, tensor tympani syndrome, and vascular lesions. Subjective tinnitus has been linked to sensorineural hearing loss caused by various deficits such as age-related hearing loss (presbycusis), noise exposure, acoustic neuroma, and Ménière’s disease (Levine, 2001), but also to other conditions such as temporomandibular joint dysfunction. Different neural mechanisms have been proposed, and tinnitus has been explained as the result of increased neural activity in the form of increased burst firing, or as a result of pathological synchronization of neural activity. Other suggested mechanisms are hypersensitivity and cortical reorganization (Rauschecker, 1999).

Prevalence and Categorization

Tinnitus is commonly a temporary sensation, and most people have experienced it. However, it may develop into a chronic condition that resists medical or surgical treatment. Prevalence figures vary slightly, but at least 10%–15% of the general population can be expected to have tinnitus. A large majority do not have severe tinnitus. Findings from epidemiological studies suggest that about 1%–3% of the adult population has severe tinnitus, in the sense that it causes marked disruption of everyday activities, mood changes, and often disrupted sleep patterns. Tinnitus has been reported in children, but in its severe form it is far more common in adults and the elderly (Davis and El Rafaie, 2000).

What distinguishes mild from severe tinnitus is not established, apart from variations in subjective ratings of intrusiveness and loudness. In particular, in attempts to determine the handicap caused by tinnitus, it has not been possible to make the determination using aspects of the tinnitus itself (e.g., loudness, character, etc.). Moreover, tinnitus has been notoriously difficult to measure objectively. However, psychological complaints are of major importance in determining the severity of tinnitus (Andersson, 2001). In line with the difficulties associated with measuring tinnitus, no consensus has been reached regarding its classification, and several schemes have been proposed. Structured interviews and validated self-report questionnaires are helpful when describing tinnitus.

Among the most influential theories on why tinnitus causes annoyance is Hallam’s psychological model of tinnitus (Hallam, Rachman, and Hinchcliffe, 1984) and Jastreboff’s (1990) neurophysiological model. The latter model puts less emphasis on conscious mechanisms involved in tinnitus perception. Basically, Jastreboff presents a conditioning model in which the tinnitus signal is conditioned to aversive reactions such as anxiety and fear.

Audiological Characteristics

Measurement of tinnitus involves subjective report and a history of its features, such as loudness, character, fluctuations, and severity. Patients have described tinnitus as tones, buzzing noises, and mixtures of buzzing and ringing. More complicated descriptions include metallic sounds and multiple tones of varying frequencies.

Since the 1930s, tinnitus has been measured by asking the patient to compare the tinnitus with an external tone or combinations of tones. Tinnitus loudness can be presented at hearing level (HL) or sensation level (SL), the latter being the level of tinnitus above hearing threshold. Further, tinnitus loudness can be matched using the tinnitus frequency (for which hearing is often impaired) or another frequency where hearing is normal (Henry and Meikle, 2000). Contralateral versus ipsilateral matching is another choice. Determining the minimal masking level is a way to quantify the intrusiveness of tinnitus by determining how loud a sound needs to be to mask the tinnitus (Henry and Meikle, 2000). Pioneering work by Feldmann (1971) revealed that tinnitus patients could be categorized according to how tones of different frequencies masked the tinnitus (so-called masking curves). For example, one type of tinnitus was equally masked by tones of low and high frequency, whereas another type of tinnitus was more easily masked by low-frequency tones. Tinnitus can often be masked by white noise, at least temporarily (Henry and Meikle, 2000).

Emotional and Cognitive Disturbances

Tinnitus patients often report difficulties with concentration, such as during reading. Until recently, few
attempts were made to measure tinnitus patients’ performance on tests of cognitive functioning, but preliminary results corroborate the self-report findings (Andersson et al., 2000).

In its severe form, tinnitus is often associated with lowered mood and depression. There are only a few studies that have endorsed structured psychiatric interviews, and most of the available data are based on questionnaires. Suicide related to tinnitus is rare. Most cases reported had associated comorbid psychiatric disturbances (Lewis, Stephens, and McKenna, 1994). Anxiety, and in particular anxious preoccupation with somatic sensations, is an important aggravating factor related to distress caused by tinnitus (Newman, Wharton, and Jacobson, 1997). Stress is often mentioned as a negative factor for tinnitus, in particular the stress of major adverse life events. Finally, sleep problems are a significant component of tinnitus patients’ complaints (McKenna, 2000).

Tinnitus and the Brain

Researchers and clinicians have long suspected that tinnitus involves certain areas of the brain, particularly those that subserve the perception and amplify the experience. Studies have been conducted on tinnitus patients’ reaction times, brainstem audiometry, evoked potentials, and magnetoencephalography. Tinnitus has recently been studied using single photon emission computed tomography, positron emission tomography (e.g., Mirz et al., 1999), and functional magnetic resonance imaging. Findings from brain imaging studies suggest that tinnitus can be objectively measured, but are not consistent. However, it is clear that tinnitus affects brain areas related to hearing and processing of sounds, but also that some involvement of the brain’s attentional and emotional systems might be involved (e.g., the amygdala).

Treatment

There is a long history of attempts to cure tinnitus. Although there are ways to alleviate the suffering, surgical and pharmacological interventions have been largely unsuccessful (Dobie, 1999). A pharmacological agent that reliably abolishes tinnitus for a short period is the local anesthetic agent lidocaine (Davies, 2001). About 60% of tinnitus sufferers respond to lidocaine administered intravenously, which in some cases totally abolishes tinnitus for a brief period. Because of side effects and the lack of effective oral analogues, lidocaine is not a viable treatment for tinnitus (Davies, 2001).

One treatment alternative for selected patients with tinnitus is antidepressants. A few studies have found positive results with respect to tinnitus annoyance, whereas more modest results were found in another study. More studies are needed, in particular to investigate the effects of selective serotonin reuptake inhibitors.

Tinnitus is rarely the only indication for surgery unless clear objective findings can identify a causal agent (e.g., nerve compression). However, when surgery is called for, the effects on tinnitus have been unclear. In some patients tinnitus disappears, but in others it remains unchanged or becomes worse (Hazell, 1990).

Alternative medicine approaches (such as Gingko biloba and acupuncture) either have not been tested or, when trials have been conducted, have yielded disappointing results (Davies, 2001).

The effects of electrical stimulation have been investigated in two forms. The first is application of electric current via transcutaneous nerve stimulation, and the second is through cochlear implantation, in which electrodes are inserted into the cochlea. The latter approach is most interesting, as cases have been reported in which tinnitus disappears while the implant is on and returns when it is turned off (Dauman, 2000).

There is a long history of attempts to treat tinnitus via maskers and, more recently, white noise generators. These are basically hearing aid-like devices that emit noise of broadband or narrow-band character. Unfortunately, there are few controlled trials on the use of masking devices or white noise generators. The studies that do exist do not support the efficacy of masking, but clinical experience suggests that they help some people (Vernon and Meikle, 2000).

More recently a treatment method called tinnitus retraining therapy (TRT) has been developed. TRT has two parts, one consisting of counseling in a directive format and the other part providing “sound enrichment” using white noise generators set at a level that does not cover tinnitus (Jastreboff and Jastreboff, 2000).

Among the treatments aimed at reducing distress, cognitive-behavioral therapy (CBT) is the most researched alternative (Andersson, 2001). CBT is a relatively brief psychological treatment approach directed at identifying and modifying maladaptive behaviors and cognitions by means of behavior change and cognitive restructuring. The focus is on applying techniques such as applied relaxation in real-life settings. There is evidence that CBT can be effective in alleviating the distress caused by tinnitus, and also that it works in a self-help format presented via the Internet (Andersson et al., 2002).

—— Gerhard Andersson

References


**Further Readings**


**Tympanometry**

Tympanometry is a measure of the acoustic admittance or ease with which acoustic energy flows into the middle ear transmission system as air pressure is varied in the ear canal. This measure is accomplished by sealing a small probe device into the ear canal. A speaker delivers a probe signal, typically 226 Hz, into the ear canal, and a microphone measures the amplitude and phase of the probe signal admitted into the middle ear system. The acoustic admittance is determined by the combined stiffness (or conversely, compliance), mass, and resistance of the eardrum and all middle ear structures. In the presence of middle ear pathology, these admittance characteristics are altered, and therefore the amplitude and phase of the probe signal measured in the ear canal are also altered. In pathology such as middle ear effusion, the eardrum is stiffened by fluid in the middle ear cavity, and only minimal acoustic energy from the probe signal is admitted into the middle ear; acoustic admittance in the plane of the eardrum is described as low. In contrast, pathology such as ossicular discontinuity makes the ear less stiff, so that most of the acoustic en-
ergy from the probe signal is admitted into the middle ear system, and acoustic admittance is high.

In addition to the loudspeaker and microphone, the probe system is connected to a pneumatic pump that adjusts ear canal pressure over a range from \(-600\) to \(+400\) daPa. The dekapascal (daPa) is the unit of pressure that has replaced mm H\(_2\)O (ANSI S3.39-1987). The two units, however, are nearly interchangeable (1 daPa = 1.02 mm H\(_2\)O).

Tympanometry became a routine clinical procedure following the landmark paper of Jerger (1970). Jerger identified three basic tympanogram shapes. A tympanogram is a graphic display of acoustic admittance measured as a function of changing ear canal pressure. A normal type A tympanogram is shown in Figure 1. Introduction of extreme pressures into the sealed ear canal stiffens the eardrum, and theoretically, all of the acoustic energy from the probe signal is reflected at the surface of the eardrum, and admittance reaches a minimum. Acoustic admittance gradually increases to a maximum, and the probe signal becomes most audible, when the pressure in the ear canal equals the pressure in the middle ear cavity. When the eustachian tube is functioning normally, atmospheric pressure of 0 daPa is maintained in the middle ear cavity, and tympanogram peak pressure (TPP) also is 0 daPa. The ear canal pressure producing peak admittance, therefore, provides an estimate of middle ear pressure. When the eardrum is retracted and negative middle ear pressure exists, the peak of the tympanograms shifts to a corresponding negative value. This tympanogram pattern is designated type C in Figure 1. The third tympanogram, designated type B, has no discernible peak and is flat. This pattern is recorded from ears with middle ear effusion (MEE), perforated eardrums, or patent pressure-equalization tubes (PET).

Tympanogram shape also has been quantified in an attempt to aid in the diagnosis of middle ear disease and to establish objective criteria for medical referral. Four commonly used calculations are depicted in Figure 2. The first, acoustic equivalent volume (V\(_{ea}\)), is an estimate of the ear canal volume between the probe device and the eardrum. This estimate typically is made using a 226-Hz probe signal and an ear canal pressure of 200 daPa. When the probe device is sealed in the ear canal, the measured acoustic admittance reflects the combined effects of the ear canal and the middle ear. Under extreme ear canal pressures, however, the eardrum theoretically becomes so stiff that acoustic admittance into the middle ear decreases to 0 mmhos. The admittance measured at extreme pressures then is attributed solely to the ear canal volume. When a 226-Hz probe signal is used, the acoustic admittance measured at 200 daPa is equal to the volume of the ear canal. In Figure 2, V\(_{ea}\) equals 0.6 cm\(^3\). In children less than 7 years old, V\(_{ea}\) ranges from 0.3 to 0.9 cm\(^3\) (Margolis and Heller, 1987; Shanks et al., 1992). In adults, V\(_{ea}\) averages 1.3 cm\(^3\) in women and 1.5 cm\(^3\) in men (Wiley et al., 1996). As a subsequent example will demonstrate, V\(_{ea}\) is useful in differentiating between intact and perforated eardrums when a flat type B tympanogram is recorded.

Peak compensated static acoustic admittance (Y\(_{tm}\)) is the amplitude of the tympanogram between the peak and 200 daPa. This measure describes the acoustic admittance of the middle ear transmission system compensated for or minus the effects of the ear canal volume. In Figure 2, Y\(_{tm}\) is 1.1 acoustic mmhos, calculated as peak admittance (1.7 mmhos) minus ear canal admittance (0.6 mmhos). Many instruments “baseline correct” at 200 daPa, so that Y\(_{tm}\) can be read directly from the y-axis. If the tympanogram in Figure 2 were baseline corrected, zero admittance would be shifted upward to correspond with the V\(_{ea}\) at 200 daPa. If a middle ear problem produces abnormally high stiffness, the

![Figure 1](image1.png)

**Figure 1.** Three patterns of tympanograms recorded using a 226 Hz probe signal. Type A is normal, type B is flat, and type C has a negative tympanogram peak pressure.

![Figure 2](image2.png)

**Figure 2.** Four calculations made on 226 Hz tympanograms: acoustic equivalent volume (V\(_{ea}\), in cm\(^3\)), peak compensated static acoustic admittance (Y\(_{tm}\), in acoustic mmhos), tympanogram width (TW, in daPa), and tympanogram peak pressure (TPP, in daPa).
amplitude of the tympanogram, or $Y_{tm}$, will decrease. Conversely, if a middle ear problem decreases the stiffness of the eardrum or middle ear, $Y_{tm}$ will increase. $Y_{tm}$ at 226 Hz normally increases slightly from infancy to adulthood, with a mean of 0.5 acoustic mmhos at 4 months to 0.7 acoustic mmhos in adulthood (Margolis and Heller, 1987; Holte, Margolis, and Cavanaugh, 1991; Roush et al., 1995; Wiley et al., 1996).

Tympanogram width (TW), defined as the width in daPa at one-half $Y_{tm}$, is a measure of the breadth of a tympanogram peak. In Figure 2, TW is 85 daPa. TW is not highly correlated with $Y_{tm}$, and therefore it provides supplemental information regarding middle ear function (Koebsell and Margolis, 1986). TW has been most useful in identifying children with middle ear effusion (MEE). In some cases of MEE, $Y_{tm}$ is normal but TW is abnormally broad. Nozza et al. (1992, 1994) reported that a TW greater than 275 daPa was associated with a high sensitivity (81%) and specificity (82%) in identifying MEE.

The fourth measure, TPP, provides an estimate of middle ear pressure or indirect measure of eustachian tube function. Figure 2 shows a normal TPP of 10 daPa. Not all individuals with negative middle ear pressure develop MEE. Results from school screening programs showed that medical referral on the basis of TPP alone resulted in unacceptably high overreferral rates, and therefore TPP is no longer used in referral criteria. A negative TPP in conjunction with a reduced $Y_{tm}$ is a much stronger indication of MEE and cause for medical referral (Feldman, 1976).

Table 1 shows means and 90% normal ranges for $V_{ea}$, $Y_{tm}$, TW, and TPP from several large-scale studies in subjects 8 weeks to 90 years old with normal middle ear transmission systems. The remaining figures depict tympanometry findings for a variety of middle ear pathologies. The probe signal frequency most commonly used to measure the admittance properties of the middle ear is 226 Hz. Although this low-frequency probe signal was selected partly at random during instrument development (Terkildsen and Scott Nielson, 1960), it remains the most commonly used probe signal. Acoustic admittance measurements at low frequencies are dominated by the stiffness characteristics of the eardrum and middle ear transmission system, whereas measurements made at high frequencies are dominated by mass characteristics. Although high-frequency probe signals of 660–1000 Hz are valuable in assessing the mass characteristics of the middle ear, the tympanogram patterns that result at high frequencies are more complex and have not enjoyed widespread use. Only low-frequency tympanograms are presented in subsequent examples, but cases where high-frequency probe signals might be advantageous are pointed out. Additional references on high-frequency tympanometry are provided.

Figure 3 shows a series of tympanograms recorded from a child recovering from a 3-month episode of otitis media with MEE. When first evaluated, the admittance tympanogram was flat (type B), with a normal $V_{ea}$ of 0.45 mmhos. Sequential pure-tone audiograms showed air–bone gaps across all frequencies, ranging from 10 dB to 55 dB, that were greatest at 250 and 4000 Hz and smallest at 2000 Hz. Over time, the tympanogram changed to a type C pattern. In early recovery, the tympanogram, shown by the heavy line, had a shallow ($Y_{tm} = 0.25$ mmhos), broad peak (TW = 200 daPa) with negative peak pressure (TPP = −100 daPa). Air–bone gaps decreased to 10–25 dB and were largest at 4000 Hz. This tympanogram pattern has been demonstrated in human temporal bones injected with middle ear fluid up to the level of the umbo, producing a mass loading effect on the eardrum (Renvall, Liden, and

---

**Table 1.** Means and 90% Ranges for $V_{ea}$, $Y_{tm}$, TW, and TPP from Several Large-Scale Studies in Subjects 8 Weeks to 90 Years Old with Normal Middle Ear Transmission Systems

<table>
<thead>
<tr>
<th>Study</th>
<th>Age (yr)</th>
<th>N</th>
<th>Statistic</th>
<th>$V_{ea}$ (cm$^3$)</th>
<th>$Y_{tm}$ (mmhos)</th>
<th>TW (daPa)</th>
<th>TPP (daPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiley et al. (1996)</td>
<td>48–90</td>
<td>2147</td>
<td>Mean</td>
<td>1.36</td>
<td>0.66</td>
<td>75</td>
<td>−23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90% range</td>
<td>0.9–2.0</td>
<td>0.2–1.5</td>
<td>35–125</td>
<td>−85 to 5</td>
</tr>
<tr>
<td>Margolis and Heller (1987)</td>
<td>19–61</td>
<td>87</td>
<td>Mean</td>
<td>1.05</td>
<td>0.78</td>
<td>77</td>
<td>−19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90% range</td>
<td>0.63–1.46</td>
<td>0.32–1.46</td>
<td>51–114</td>
<td>−83 to 0</td>
</tr>
<tr>
<td>Nozza et al. (1992, 1994)</td>
<td>3–16</td>
<td>130</td>
<td>Mean</td>
<td>0.74</td>
<td>0.55</td>
<td>100</td>
<td>−30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90% range</td>
<td>0.42–0.97</td>
<td>0.22–0.92</td>
<td>59–151</td>
<td>−139 to 11</td>
</tr>
<tr>
<td>Roush et al. (1995)</td>
<td>0.5–2.5</td>
<td>+1636</td>
<td>Mean</td>
<td>0.90</td>
<td>0.78</td>
<td>104</td>
<td>−34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90% range</td>
<td>0.60–1.35</td>
<td>0.40–1.39</td>
<td>60–168</td>
<td>−207 to 15</td>
</tr>
<tr>
<td>Shanks et al. (1992)</td>
<td>8 wk–7 yr</td>
<td>334</td>
<td>Mean</td>
<td>0.58</td>
<td>0.45</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90% range</td>
<td>0.20–0.70</td>
<td>0.23</td>
<td>102–204</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** $V_{ea}$, acoustic equivalent volume; $Y_{tm}$, peak compensated static acoustic admittance; TW, tympanogram width; TPP, tympanogram peak pressure.
The mass effect is greatest at high frequencies, as reflected by large air–bone gaps at 4000 Hz, and is accentuated when tympanometry is performed using high-frequency probe signals such as 600–800 Hz. Further resolution of the otitis media produced a type C tympanogram, increasing to normal $Y_{tm}$ (0.35 mmhos) with a TPP of $-200$ daPa. Small air–bone gaps of 15–20 dB at this time were confined to the 250–1000 Hz range and were virtually closed at 4000 Hz, indicating increased stiffness of the eardrum from the negative middle ear pressure without the mass loading effects of the middle ear fluid. The tympanogram gradually returned to a type A. This case study demonstrates a variety of tympanogram patterns associated with otitis media. Rather than being a drawback, the various tympanogram shapes help clinicians track the resolution of MEE.

The American Speech-Language-Hearing Association (1997) has developed guidelines for screening infants and children for chronic middle ear disorders with the potential for causing significant hearing loss or long-lasting speech, language, and learning deficits. Medical referral is advised for infants when $Y_{tm}$ is less than 0.2 mmhos or TW is greater than 235 daPa, and for 1- to 5-year-olds when $Y_{tm}$ is less than 0.3 mmhos or TW is greater than 200 daPa if these abnormal findings persist at a 6–8-week rescreening. Immediate medical referral is recommended for otalgia, otorrhea, or eardrum perforation noted otoscopically or from a flat tympanogram with $V_{ea}$ greater than 1.0 cm$^3$. Screening guidelines are not available for infants less than 7 months old. In this age group, tympanogram shapes at 226 Hz are irregular and difficult to interpret (Holte, Margolis, and Cavanaugh, 1991).

Figure 4 displays three type B tympanograms. The bottom tympanogram was recorded from an ear with an intact eardrum and MEE; $V_{ea}$ of 0.45 cm$^3$ is normal for a child’s ear canal. The other two tympanograms also are flat but $V_{ea}$ is 3.25 cm$^3$ in one case and greater than 5.0 cm$^3$ in the other. The middle tympanogram was recorded from an ear with a patent PET, and the top tympanogram was recorded from an ear with a traumatic perforation from a Q-tip. $V_{ea}$ often is larger with a traumatic eardrum perforation than with a perforation associated with chronic middle ear disease and poorly developed mastoid air-cell system (Andreasson, 1977).

In a child less than 7 years old, a volume greater than 1.0 cm$^3$ is indicative of a perforated eardrum, whereas in adults the volume must exceed 2.5 cm$^3$ (Shanks et al., 1992). Volumes exceeding these ranges clearly indicate a perforated eardrum, but flat tympanograms with smaller volumes do not necessarily rule out eardrum perforation. A flat tympanogram with a normal $V_{ea}$ can also be recorded from an ear with eardrum perforation and cholesteatoma filling the middle ear space and closing off the mastoid air-cell system. Case history and otoscopic examination are very important in these cases. No consistent pattern of hearing loss is associated with eardrum perforation; air–bone gaps can be absent or as large as 50–70 dB if necrosis of the incus also occurs.

Figure 5 demonstrates that otosclerosis also is associated with a variety of tympanogram shapes. Tympanograms vary from a normal type A pattern (shown in Fig. 1) to a low-admittance, stiff pattern (type $A_s$) shown by the lower tympanogram in Figure 5. A third pattern frequently recorded in otosclerosis is a normal type A pattern, but with a narrow tympanogram width (Shanks, 1984; Shahnaz and Polka, 1997). Pure-tone audiometry in otosclerosis shows a stiffness tilt, with the largest air–bone gaps at low frequencies and the smallest air–bone gap near 2000 Hz. Otosclerosis is virtually the only middle ear pathology where significant air–bone gaps are measured in conjunction with a normal type A tympanogram.

Figure 6 shows two cases of type A tympanograms with abnormally high $Y_{tm}$ (2.5 mmhos). Normal tympanograms with deep peaks sometimes are designated type $A_d$. The bottom tympanogram was recorded from
an ear with traumatic ossicular discontinuity; the audiogram showed a maximum conductive hearing loss with 30–70 dB air–bone gaps. The top tympanogram was recorded from an ear with a monomeric eardrum resulting from a healed perforation; the audiogram showed slight air–bone gaps at only 3000 and 4000 Hz. This $A_d$ pattern also is typical of ears with tympanosclerotic plaques on the eardrum and status post stapedectomy. These cases of high-admittance pathology are another indication for high-frequency tympanometry. High-frequency tympanograms in ears with ossicular discontinuity typically exhibit broader, more undulating peaks than tympanograms recorded from ears with eardrum pathology. In cases of high $Y_{im}$, otoscopic examination of the eardrum is crucial; high-admittance pathology of the eardrum can dominate or mask low-admittance pathology such as otosclerosis. Pure-tone audiometry also is invaluable in these cases. Eardrum pathology alone does not produce a significant conductive hearing loss, whereas ossicular discontinuity results in a maximum conductive hearing loss.

The preceding cases demonstrate that tympanometry is most beneficial when used as one of a battery of tests that also include case history, otoscopic examination, and pure-tone audiometry. The cases also demonstrate that each unique middle ear problem does not produce one and only one tympanogram pattern. On the contrary, a single pathology can produce several different tympanometry patterns, and conversely, a single tympanogram pattern can result from several different middle ear problems. When used with a battery of tests, however, the contribution from tympanometry can be unique and informative.

See also MIDDLE EAR ASSESSMENT IN THE CHILD.

—Janet E. Shanks

References


**Further Readings**


**Vestibular Rehabilitation**

The concept of prescribing exercise for persons with dizziness was first described by Cooksey and Cawthorne in the 1950s (Cawthorne, 1944; Cooksey, 1946). Today, exercise for persons with vestibular disorders is considered to be the standard of care (Cowand et al., 1998; Herdman, 1990; Herdman, 1992; Herdman et al., 1995; Herdman and Whitney, 2000). Exercises are specifically prescribed that help the person with a vestibular disorder either compensate for or adapt to the impairment (Shepard and Telian, 1993). Knowledge of vestibular anatomy, physiology, pathologies involved, and an in-depth understanding of how various interventions can affect outcome is very important for effective treatment of persons with vestibular disorders. Exercises to decrease the risk of falling, improve balance and postural control, improve confidence, and decrease the subjective feelings of dizziness also seem to decrease a patient’s anxiety (Jacob et al., 2000). Vestibular exercise programs have been shown to enhance the speed and degree of recovery (Herdman et al., 1995; Horak et al., 1992; Krebs et al., 1993; Strupp et al., 1998; Yardley et al., 1998).

Common conditions often referred for vestibular physical and occupational therapy include benign paroxysmal positional vertigo (BPPV) (Blakely, 1994; Herdman et al., 1993; Lynn et al., 1995), bilateral vestibulopathy (Brown et al., 2001; Krebs et al., 1993; Telian et al., 1991), endolymphatic hydrops, labyrinthine concussion (Cowand et al., 1998; Fujino et al., 1996; Horak et al., 1992; Shepard et al., 1990; Shepard et al., 1993; Smith-Wheelock et al., 1991), labyrinthitis (Cowand et al., 1998; Fujino et al., 1996; Shepard et al., 1990; Shepard et al., 1993; Smith-Wheelock et al., 1991), Ménière’s disease (Cowand et al., 1998; Fujino et al., 1996; Shepard et al., 1990; Shepard et al., 1993; Smith-Wheelock et al., 1991), perilymph fistula, and vestibular neuritis. Central diagnoses include cervicogenic dizziness, brainstem hemorrhage (Cowand et al., 1998; Horak et al., 1992; Shepard et al., 1990; Shepard et al., 1995; Smith-Wheelock et al., 1991), posttraumatic anxiety symptoms, stroke/transient ischemic attacks (TIA), traumatic head injury (Cowand et al., 1998; Horak et al., 1992; Shepard et al., 1990; Shepard et al., 1993), and migraine-related vestibulopathy (Cass et al., 1997; Whitney et al., 2000). Psychiatric disorders that have been reported to manifest with a component of dizziness include panic disorders (Jacob et al., 2000), agoraphobia (Jacob et al., 2000), and hyperventilation syndrome. The most common nonvestibular causes of dizziness are low blood pressure and medication-induced dizziness (Furman and Whitney, 2000).

Persons with vestibular disorder present with various complaints, and often report experiencing balance dysfunction, dizziness, vertigo, anxiety about their symptoms, space and motion complaints (symptoms elicited by a specific visual stimulus pattern (Furman and Cass, 1996), and fear of falling. They may describe visual disturbances, disequilibrium, and dizziness occurring while they are at work, at home, or engaged in leisure activities. Common visual problems experienced include difficulty focusing while reading, “bouncing” of the visual world as they move (oscillopsia), impaired smooth pursuit, saccades, and vergence. Balance problems frequently noted include increased sway while standing, an inability to stand still, walking with a wide-based gait, veering during walking, adduction or crossing their legs during gait, difficulty walking in the dark or on uneven surfaces, bumping into things, or falling.

Tinnitus, difficulty hearing, and aural fullness are related cochlear signs reported by persons with vestibular
disorders. Descriptions of problems related to the head include dizziness, spinning, headache, pressure, neck pain, a swimming sensation, and heaviness. Often, persons with vestibular disorders report fatigue and difficulty concentrating. All of these problems contribute to making vestibular disorders difficult to treat, as affected individuals may have multiple symptoms and frequently have more than one diagnosis.

A physical therapy evaluation provides information on impairments and functional deficits so that appropriate intervention can be determined. A thorough workup by the physician and vestibular function tests help direct the physical therapy evaluation and intervention. The patient history should include goals of treatment, premorbid health, current and premorbid activity level, and a description of the onset, frequency, duration, and severity of the dizziness and imbalance. Not all persons with vestibular disorders experience both dizziness and imbalance. Identifying the positions or situations that exacerbate or relieve the symptoms can afford valuable insight into the cause of the problem. Gaining an understanding of the magnitude of the functional deficits is very important. The intensity and duration of symptoms, the degree to which symptoms impede activities of daily living, and how symptoms affect social activities help to determine intervention.

A thorough exploration of the individual’s history of falling can also provide insight into the physiology of the condition and the necessary treatment (Herdman et al., 2000; Whitney, Hudak, and Marchetti, 2000). Not simply the number of falls but also the conditions of the fall, the frequency of falling, and whether medical treatment was necessary are all important in the assessment of the person with a vestibular disorder. Fear of falling in individuals who have fallen may constrict their willingness to move (Tinetti and Powell, 1993).

The patient’s medical and surgical history will affect the prescription of an exercise program. Persons with premorbid orthopedic and cardiac limitations need to be carefully monitored to ensure that they are safe with the exercise program. Frail, older adults may need to be seen more frequently in order to ensure compliance and safety with their exercises.

Typically, the range of motion of the joints, muscle strength, sensation, vision, motions that provoke symptoms, balance, and gait are all determined before an exercise program is started. Because of the influence of somatosensation on balance, it is important to assess range of motion and sensation, particularly in the ankles and cervical region. The visual assessment includes testing the function of the ocular muscles, including saccades and smooth pursuit, as well as the function of the vestibular ocular reflex.

Quantification of the movements and positions that trigger symptoms of dizziness not only provides information on the cause of the symptoms but may also help in selecting activities for treatment. Therapists commonly ask patients to move into and out of supine and side-lying positions, and then have the patients rate their symptoms on a verbal analogue scale and indicate the duration of the symptoms (Norre and De Weerdt, 1980; Smith-Wheelock et al., 1991). Monitoring the intensity and duration of the symptoms over the length of treatment can provide information on the recovery of the patient.

Two aspects of postural control should be evaluated, the ability to move the center of gravity within the base of support and the ability to utilize available sensory information for balance. The ability to move the center of gravity within the base of support is determined by asking the person to perform tasks such as shifting his or her weight while standing and then reaching for objects. This can be quantified by measuring how far the person can reach (i.e., the Functional Reach test [Duncan et al., 1990] or the multidirectional Reach test [Newton, 2001]) or how long they can maintain a position (i.e., standing on one foot or in tandem). In addition, the clinician can ask patients to twist their trunk, pick up objects from different surfaces, or stand on a narrow base of support to determine how stable they are doing functional activities. Having the person with a vestibular disorder stand on high-density foam with the eyes open and then closed (Clinical Test of Sensory Interaction and Balance, or CTSIB) (Shumway-Cook and Horak, 1986) can help the therapist determine the fall risk (Anacker and Di Fabio, 1992), and also how well the patient uses the sensory information that he or she has available (Shumway-Cook and Horak, 1986). Scores on the CTSIB have been shown to correlate with conditions 4 and 5 of computed dynamic posturography (Anacker and Di Fabio, 1992; Weber and Cass, 1993).

Patients with vestibular disorders often describe difficulty walking, especially under varying sensory conditions such as walking with head turns, in dimmed or absent lighting, or with movement in the environment. Assessment of the person’s gait during various functional tasks and under different sensory conditions is crucial. Persons with a vestibular disorder are asked to walk, walk at different speeds, walk over and around objects, and walk with various head movements in order to determine how stable they are during ambulation (Whitney et al., 2000a; Whitney and Herdman, 2000). These tasks can be quantified using the time it takes to complete a task or by qualitatively describing the movements, as in the Dynamic Gait Index (Shumway-Cook and Woollacott, 1995).

The goals of vestibular rehabilitation include decreasing the risk of falling, improving gait stability, improving the person’s dynamic and static postural control, decreasing symptoms, and enhancing the individual’s ability to carry out activities of daily living and to work. The achievement of these goals is through exercise and the practice of activities in a safe environment. Customized exercise programs are better than exercise handouts provided without direction as to which exercises are most important to perform (Shepard and Telian, 1995).

Vestibular rehabilitation intervention is prescribed individually for each patient. For patients with periph-
erel vestibular lesions, vestibular rehabilitation exercises are thought to promote compensation or recalibration of the vestibular system, specifically the vestibulo-ocular reflex (VOR). The system appears to recalibrate because an error signal is created from the slip of an image on the retina (Robinson, 1976; Fetter and Zee, 1988). The use of eye and head movements as an exercise to change the gain of the VOR results in a change in the inhibition of the activity in the vestibular nuclei, and consequently in enhanced patient function. Recovery of the VOR is frequency specific. Stimulation of the VOR through exercises must be performed at varying frequencies for maximal functional recovery (Godaux, Halleux, and Gobert, 1983; Lisberger, Miles, and Optican, 1983).

Activity after a lesion to the vestibular system is important. Animals that moved freely after surgery had faster functional recovery (Lacour, Roll, and Appaix, 1976). Patients with vestibular disorders had faster recovery and improved function when they increased their activity early after surgery (Herdman et al., 1995).

Some persons with vestibular disorders have little or no remaining vestibular function, owing to disease or ototoxicity. These persons must learn how to use remaining sensory function such as somatosensation and vision. In addition, receptors in the neck can assist in stabilizing vision and posture, although in patients with intact vestibular systems, the cervical ocular reflex contributes little to gaze stability. The cervical ocular reflex performs maximally at lower frequencies and is accentuated in patients with bilateral vestibular loss (Kasai and Zee, 1978; Bronstein and Hood, 1986). Smooth pursuits and saccades can also assist in stabilizing vision at slow speeds (Kasai and Zee, 1978; Segal and Katsar-Kas, 1988; Leigh et al., 1994). Patients who have no function in the vestibular system will never be able to walk in the dark, will have great difficulty walking on uneven surfaces, and will never be able to read and walk at the same time. Driving a car without any vestibular function is impossible because the visual field jumps (oscillopsia), especially as the car goes over bumps.

Patients with balance disorders are taught to maximize the sensory function that remains, to substitute for sensory loss, and to predetermine when they will have difficulty with balance so that they can modify their behavior. Exercises are prescribed to enhance the use of vestibular, visual, and somatosensory inputs. Habituation exercises may be recommended for patients with dizziness provoked by specific position changes.

The outcomes of vestibular rehabilitation have included a decrease in dizziness and vertigo, a decrease in the number of falls, improved gait, decreased neck pain, improved VOR, greater balance confidence, decreased anxiety, improvements in activities of daily living, and improvements in the perceived disability. Generally, persons with peripheral vestibular disorders have a better prognosis than those with central vestibular disorders. Persons with both central and peripheral vestibular disorders have a poorer prognosis. All persons with vestibular disorders should have an opportunity to work with a knowledgeable physical or occupational therapist, because quality of life can be improved with vestibular rehabilitation.

—Susan L. Whitney and Diane M. Wrisley

References


**Further Readings**


Contributors

General Editor

Raymond D. Kent
University of Wisconsin-Madison, Madison, Wisconsin

Advisory Editors

Voice Disorders in Children
Steven D. Gray (deceased)
University of Utah, Salt Lake City, Utah

Robert E. Hillman
Massachusetts Eye and Ear Infirmary, Boston, Massachusetts

Voice Disorders in Adults

Lawrence D. Shriberg
University of Wisconsin-Madison, Madison, Wisconsin

Speech Disorders in Children

Joseph R. Duffy
Mayo Clinic, Rochester, Minnesota

Speech Disorders in Adults

Mabel L. Rice
University of Kansas, Lawrence, Kansas

David A. Swinney
University of California at San Diego, San Diego, California

and

Lewis P. Shapiro
San Diego State University, San Diego, California

Language Disorders in Children

Sandra Gordon-Salant
University of Maryland, College Park, Maryland

Language Disorders in Adults

Hearing Disorders in Children

Pelagie M. Beeson
University of Arizona, Tuscon, Arizona
AGRAPHIA

Gerald S. Berke
University of California at Los Angeles School of Medicine, Los Angeles, California LARYNGEAL REINNERVATION PROCEDURES

Ken Bleile
University of Northern Iowa, Cedar Falls, Iowa SPEECH DISORDERS IN CHILDREN: BIRTH-RELATED RISK FACTORS

Joel H. Blumin
University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania LARYNGEAL REINNERVATION PROCEDURES

Sheila E. Blumstein
Brown University, Providence, Rhode Island PHONOLOGY AND ADULT APHASIA

Carol A. Boliek
University of Alberta, Edmonton, Alberta SPEECH DEVELOPMENT IN INFANTS AND YOUNG CHILDREN WITH A TRACHEOSTOMY

Louis D. Braida
Massachusetts Institute of Technology, Cambridge, Massachusetts FREQUENCY COMPRESSION

Mitchell F. Brin
Columbia University, New York, New York HYPOKINETIC LARYNGEAL MOVEMENT DISORDERS

Contributing Authors

Faith W. Akin
James H. Quillen VA Medical Center, Mountain Home, Tennessee ELEKTRONYSTAGMOGRAPHY

Jont B. Allen
University of Illinois at Urbana-Champaign, Beckman Institute, Urbana, Illinois AMPLITUDE COMPRESSION IN HEARING AIDS

Gerhard Andersson
Uppsala University, Uppsala, Sweden TINNITUS

Daniel H. Ashmead
Vanderbilt University Medical Center, Nashville, Tennessee AUDITION IN CHILDREN, DEVELOPMENT OF

Laura J. Ball
University of Nebraska Medical Center, Omaha, Nebraska AUGMENTATIVE AND ALTERNATIVE COMMUNICATION APPROACHES IN ADULTS AUGMENTATIVE AND ALTERNATIVE COMMUNICATION APPROACHES IN CHILDREN

Martin J. Ball
University of Louisiana at Lafayette, Lafayette, Louisiana SPEECH ASSESSMENT, INSTRUMENTAL

Shari R. Baum
McGill University, Montreal, Quebec PROSODIC DEFICITS

Kathryn A. Bayles
University of Arizona, Tuscon, Arizona DEMENTIA
Bonnie Brinton
Brigham Young University, Provo, Utah
PRAGMATIC
PSYCHOSOCIAL PROBLEMS ASSOCIATED WITH COMMUNICATIVE DISORDERS

Hiram Brownell
Boston College, Chestnut Hill, Massachusetts
RIGHT HEMISPHERE LANGUAGE AND COMMUNICATION FUNCTIONS IN ADULTS

Hugh W. Buckingham
Louisiana State University, Baton Rouge, Louisiana
PHONOLOGICAL ANALYSIS OF LANGUAGE DISORDERS IN APHASIA

Eugene H. Buder
University of Memphis, Memphis, Tennessee
ACOUSTIC ASSESSMENT OF VOICE

Angela Burda
University of Northern Iowa, Cedar Falls, Iowa
SPEECH DISORDERS IN CHILDREN: BIRTH-RELATED RISK FACTORS

Thomas Campbell
University of Pittsburgh and Children’s Hospital of Pittsburgh, Pittsburgh, Pennsylvania
EARLY RECURRENT OTITIS MEDIA AND SPEECH DEVELOPMENT

David Caplan
Massachusetts General Hospital, Boston, Massachusetts
APHASIC SYNDROMES: CONNECTIONIST MODELS

Janina K. Casper
SUNY-University Medical Center, Syracuse, New York
VOCAL HYGIENE

Hugh W. Catts
University of Kansas, Lawrence, Kansas
LANGUAGE IMPAIRMENT AND READING DISABILITY

Michael R. Chial
University of Wisconsin-Madison, Madison, Wisconsin
HEARING PROTECTION DEVICES

Li-Rong Lilly Cheng
San Diego State University, San Diego, California
SPEECH AND LANGUAGE ISSUES IN CHILDREN FROM ASIAN-PACIFIC BACKGROUNDS

Melissa Cheslock
Georgia State University, Atlanta, Georgia
AUGMENTATIVE AND ALTERNATIVE COMMUNICATION: GENERAL ISSUES

Chris Code
University of Exeter, Exeter, United Kingdom
APHASIA TREATMENT: PSYCHOSOCIAL ISSUES

Rohyn M. Cox
University of Memphis, Memphis, Tennessee
HEARING AID FITTING: EVALUATION OF OUTCOMES

Martha Crago
McGill University, Montreal, Quebec
LANGUAGE IMPAIRMENT IN CHILDREN: CROSS-LINGUISTIC STUDIES

Carl C. Crandell
University of Florida, Gainesville, Florida
CLASSROOM ACOUSTICS

Richard F. Curlee
University of Arizona, Tuscon, Arizona
STUTTERING

Anne Cutler
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
SEGMENTATION OF SPOKEN LANGUAGE BY NORMAL ADULT LISTENERS

Barbara L. Davis
University of Texas at Austin, Austin, Texas
DEVELOPMENTAL APRAXIA OF SPEECH

Peter A. de Villiers
Smith College, Northampton, Massachusetts
LANGUAGE OF THE DEAF: ACQUISITION OF ENGLISH

Allan O. Diefendorf
Indiana University School of Medicine, Indianapolis, Indiana
PEDiatric AUDIOLoGY: THE TEST BATTERY APPROACH

Harvey Dillon
The National Acoustic Laboratories, Chatswood, Australia
HEARING AIDS: PRESCRIPTIVE FITTING

Christine Dollaghan
University of Pittsburgh, Pittsburgh, Pennsylvania
EARLY RECURRENT OTITIS MEDIA AND SPEECH DEVELOPMENT

Patrick J. Doyle
VA Pittsburgh Healthcare System and University of Pittsburgh, Pittsburgh, Pennsylvania
APRAXIA OF SPEECH: NATURE AND PHENOMENOLOGY

Philip C. Doyle
University of Western Ontario, London, Ontario
LARYNGEAL VOICE AND SPEECH REHABILITATION

Judy R. Dubno
Medical University of South Carolina, Charleston, South Carolina
PRESBYACUSIS

Joseph R. Duffy
Mayo Clinic, Rochester, Minnesota
DYSARTHRIAS: CHARACTERISTICS AND CLASSIFICATION

Anh Duong
University of Montreal, Montreal, Quebec
DISCOURSE IMPAIRMENTS

Tanya L. Eadie
University of Western Ontario, London, Ontario
LARYNGEAL VOICE AND SPEECH REHABILITATION

David A. Eddins
SUNY-Buffalo, Buffalo, New York
TEMPORAL RESOLUTION

Mary Louise Edwards
Syracuse University, Syracuse, New York
PHONETIC TRANSCRIPTIONS OF CHILDREN’S SPEECH

Wiltrud Fassbinder
University of Pittsburgh, Pittsburgh, Pennsylvania
RIGHT HEMISPHERE LANGUAGE DISORDERS

Heidi M. Feldman
University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania
LANGUAGE DEVELOPMENT IN CHILDREN WITH FOCAL LESIONS
John A. Ferraro
University of Kansas Medical Center, Kansas City, Kansas
ELECTROCOCHLEOGRAPHY

Marc E. Fey
University of Kansas Medical Center, Kansas City, Kansas
PRESCCHOOL LANGUAGE INTERVENTION

W. Tecumseh Fitch
Harvard University, Cambridge, Massachusetts
VOCAL PRODUCTION SYSTEM: EVOLUTION

Cynthia G. Fowler
University of Wisconsin-Madison, Madison, Wisconsin
AUDITORY BRAINSTEM RESPONSE IN ADULTS

Cynthia Fox
University of Colorado, Boulder, Colorado
HYPOKINETIC LARYNGEAL MOVEMENT DISORDERS

Naama Friedmann
Tel Aviv University, Tel Aviv, Israel
SYNTACTIC TREE PRUNING

Martin Fujiki
Brigham Young University, Provo, Utah
PRAGMATICS
PSYCHOSOCIAL PROBLEMS ASSOCIATED WITH COMMUNICATIVE DISORDERS

Fred Genesee
McGill University, Montreal, Quebec
BILINGUALISM AND LANGUAGE IMPAIRMENT

LouAnn Gerken
University of Arizona, Tucson, Arizona
LINGUISTIC ASPECTS OF CHILD LANGUAGE IMPAIRMENT—PROSODY

Bruce Gerratt
University of California at Los Angeles School of Medicine, Los Angeles, California
VOICE QUALITY, PERCEPTUAL EVALUATION

Judith A. Gierut
Indiana University, Bloomington, Indiana
SPEECH ASSESSMENT IN CHILDREN: DESCRIPTIVE LINGUISTIC METHODS

Lisa Go¤man
Purdue University, West Lafayette, Indiana
MOTOR SPEECH INVOLVEMENT IN CHILDREN

Brian Goldstein
Temple University, Philadelphia, Pennsylvania
SPEECH ISSUES IN CHILDREN FROM LATINO BACKGROUNDS

Michael P. Gorga
Boys Town National Research Hospital, Omaha, Nebraska
OTOACOUSTIC EMISSIONS IN CHILDREN

Yosef Grodzinsky
McGill University, Montreal, Quebec; Tel Aviv University, Tel Aviv, Israel
TRACE DELETION HYPOTHESIS

Murray Grossman
University of Pennsylvania, Philadelphia, Pennsylvania
ALZHEIMER’S DISEASE

Timothy C. Hain
Northwestern University School of Medicine, Chicago, Illinois
OTOOTOXIC MEDICATIONS

Joseph W. Hall
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
HEARING LOSS AND THE MASKING-LEVEL DIFFERENCE

Helen M. Hanson
Massachusetts Institute of Technology, Cambridge, Massachusetts
VOICE ACOUSTICS

M. N. Hegde
California State University at Fresno, Fresno, California
SPEECH DISORDERS IN CHILDREN: BEHAVIORAL APPROACHES TO REMEDIATION

Nancy Helm-Estabrooks
Boston University School of Medicine, Boston, Massachusetts
PERSEVERATION

Janet Helminski
Midwestern University School of Physical Therapy, Downers Grove, Illinois
OTOTOXIC MEDICATIONS

Gregory Hickok
University of California at Irvine, Irvine, California
AUDITORY-MOTOR INTERACTION IN SPEECH AND LANGUAGE
FUNCTIONAL BRAIN IMAGING

Argye E. Hillis
Johns Hopkins University School of Medicine, Baltimore, Maryland
ALEXIA

Robert E. Hillman
Massachusetts Eye and Ear Infirmary and Massachusetts General Hospital Institute of Health Professions, Boston, Massachusetts; Harvard Medical School, Cambridge, Massachusetts
AERODYNAMIC ASSESSMENT OF VOCAL FUNCTION

Jacqueline J. Hinckley
University of South Florida, Tampa, Florida
COMMUNICATION DISORDERS IN ADULTS: FUNCTIONAL APPROACHES TO APHASIA

Megan M. Hodge
University of Alberta, Edmonton, Alberta
SPEECH DISORDERS IN CHILDREN: MOTOR SPEECH DISORDERS OF KNOWN ORIGIN

Barbara Hodson
Wichita State University, Wichita, Kansas
PHONOLOGICAL AWARENESS INTERVENTION FOR CHILDREN WITH EXPRESSIVE PHONOLOGICAL IMPAIRMENTS

Erica Ho¤
Florida Atlantic University, Davie, Florida
POVERTY: EFFECTS ON LANGUAGE

Jeanette D. Hoit
University of Arizona, Tucson, Arizona
VENTILATOR-SUPPORTED SPEECH PRODUCTION

Audrey L. Holland
University of Arizona, Tucson, Arizona
COMMUNICATION DISORDERS IN ADULTS: FUNCTIONAL APPROACHES TO APHASIA

Linda J. Hood
Kresge Hearing Research Laboratory, New Orleans, Louisiana
GENETICS AND CRANIOFACIAL ANOMALIES

Aquiles Iglesias
Temple University, Philadelphia, Pennsylvania
LANGUAGE DISORDERS IN LATINO CHILDREN

David Ingram
Arizona State University, Tempe, Arizona
SPEECH DISORDERS IN CHILDREN: CROSS-LINGUISTIC DATA
Yves Joanette  
University of Montreal, Montreal, Quebec  
DISCOURSE IMPAIRMENTS

Joel C. Kahane  
University of Memphis, Memphis, Tennessee  
ANATOMY OF THE HUMAN LARYNX

Richard C. Katz  
Carl T. Hayden VA Medical Center, Phoenix, Arizona;  
Arizona State University, Tuscon, Arizona  
APHASIA TREATMENT: COMPUTER-AIDED REHABILITATION

Gitte Keidser  
The National Acoustic Laboratories, Chatswood, Australia  
HEARING AIDS: PRESCRIPTIVE FITTING

Ray D. Kent  
University of Wisconsin-Madison, Madison, Wisconsin  
INSTRUMENTAL ASSESSMENT OF CHILDREN’S VOICE

Herman Kolk  
Catholic University of Nijmegen, Nijmegen, The Netherlands  
AGRAMMATISM

Jody Kreiman  
University of California at Los Angeles School of Medicine,  
Los Angeles, California  
VOICE QUALITY, PERCEPTUAL EVALUATION OF

Charissa R. Lansing  
University of Illinois at Champaign-Urbana, Champaign,  
Illinois  
SPEECHREADING TRAINING AND VISUAL TRACKING

Charles R. Larson  
Northwestern University, Evanston, Illinois  
VOCALIZATION, NEURAL MECHANISMS

David P. Lau  
Singapore General Hospital, Singapore  
INFECTION DISEASES AND INFLAMMATORY CONDITIONS OF THE LARYNX

Matti Lehtihalmes  
University of Oulu, Oulu, Finland  
APHASIA, WERNICKE’S

Laurence B. Leonard  
Purdue University, West Lafayette, Indiana  
SPECIFIC LANGUAGE IMPAIRMENT IN CHILDREN

Barbara A. Lewis  
Rainbow Babies and Children’s Hospital and Case Western  
Reserve University, Cleveland, Ohio  
SPEECH DISORDERS: GENETIC TRANSMISSION

Marcia C. Linebarger  
Moss Rehabilitation Research Institute and Psycholinguistic  
REVERSABILITY/MAPPING DISORDERS

Sue Ellen Linville  
Marquette University, Milwaukee, Wisconsin  
VOICE DISORDERS OF AGING

Diane Frome Loeb  
University of Kansas, Lawrence, Kansas  
COMMUNICATION DISORDERS IN INFANTS AND TODDLERS

Brenda L. Lonsbury-Martin  
University of Colorado, Denver, Colorado  
OTOACOUSTIC EMISSIONS

Christy L. Ludlow  
National Institute of Neurological Disorders and Stroke and  
National Institutes of Health, Bethesda, Maryland  
LARYNGEAL MOVEMENT DISORDERS: TREATMENT WITH  
BOTULINUM TOXIN

Robert C. Marshall  
University of Kentucky, Lexington, Kentucky  
SPEECH DISORDERS IN ADULTS, PSYCHOGENIC

Julie J. Masterson  
Southwestern Missouri State University, Springfield, Missouri  
SPEECH AND LANGUAGE DISORDERS IN CHILDREN: COMPUTER-BASED APPROACHES

Patricia McCarthy  
Rush University and Rush-Presbyterian-St. Luke’s Medical  
Center, Chicago, Illinois  
FUNCTIONAL HEARING LOSS IN CHILDREN

Rebecca J. McCauley  
University of Vermont, Burlington, Vermont  
SPEECH SOUND DISORDERS IN CHILDREN: DESCRIPTION AND  
CLASSIFICATION

Karla McGregor  
Northwestern University, Evanston, Illinois  
SEMANTICS

Malcolm R. McNeil  
University of Pittsburgh, Pittsburgh, Pennsylvania  
APRAXIA OF SPEECH: NATURE AND PHENOMENOLOGY  
APRAXIA OF SPEECH: TREATMENT  
ATTENTION AND LANGUAGE

Lise Menn  
University of Colorado, Boulder, Colorado  
APHASIOLOGY, COMPARATIVE

Jon F. Miller  
University of Wisconsin-Madison, Madison, Wisconsin  
COMMUNICATION SKILLS OF PEOPLE WITH DOWN SYNDROME

John H. Mills  
Medical University of South Carolina, Charleston, South  
Carolina  
NOISE-INDUCED HEARING LOSS  
PRESBYACUSIS

Murray D. Morrison  
University of British Columbia, Vancouver, British Columbia  
INFECTION DISEASES AND INFLAMMATORY CONDITIONS OF THE LARYNX

Dave J. Muller  
University College Suffolk, Ipswich, United Kingdom  
APHASIA TREATMENT: PSYCHOSOCIAL ISSUES

Bruce E. Murdoch  
University of Queensland, Brisbane, Australia  
LANGUAGE DISORDERS IN ADULTS: SUBCORTICAL INVOLVEMENT

Thomas Murry  
Columbia University, New York, New York  
ASSESSMENT OF FUNCTIONAL IMPACT OF VOICE DISORDERS  
VOICE THERAPY FOR NEUROLOGICAL AGING-RELATED VOICE  
DISORDERS

Penelope S. Myers  
Mayo Clinic, Rochester, Minnesota  
APROSODIA

Igor V. Náblek  
University of Tennessee, Knoxville, Tennessee  
TEMPORAL INTEGRATION

Stephen T. Neely  
Boys Town National Research Hospital, Omaha, Nebraska  
OTOACOUSTIC EMISSIONS IN CHILDREN
Arlene C. Neuman  
CUNY-The Graduate Center, New York, New York  
HEARING AIDS: SOUND QUALITY

Marilyn A. Nippold  
University of Oregon, Eugene, Oregon  
LANGUAGE DISORDERS IN SCHOOL-AGE CHILDREN: ASPECTS OF ASSESSMENT

Robert J. Nozza  
Temple University School of Medicine, Philadelphia, Pennsylvania  
MIDDLE EAR ASSESSMENT IN THE CHILD

Janna B. Oetting  
Louisiana State University, Baton Rouge, Louisiana  
DIALECT SPEAKERS  
DIALECT VERSUS DISORDER

Jennifer Ogar  
VA Northern California Health Care System, Martinez, California  
APHASIA: THE CLASSICAL SYNDROMES

Gloria Streit Olness  
University of Texas at Dallas, Dallas, Texas  
DISCOURSE

Robert F. Orlikoff  
CUNY-Hunter College, New York, New York  
ELECTROGLOTTOGRAPHIC ASSESSMENT OF VOICE

Mary Joe Osberger  
Advanced Bionics Corporation, Valencia, California  
COCHLEAR IMPLANTS  
COCHLEAR IMPLANTS IN CHILDREN

Johanne Paradis  
University of Alberta, Edmonton, Alberta  
LANGUAGE IMPAIRMENT IN CHILDREN: CROSS-LINGUISTIC STUDIES

Rhea Paul  
Southern Connecticut State University and Yale Child Study Center, New Haven, Connecticut  
AUTISM

Diane Paul-Brown  
American Speech-Language-Hearing Association, Rockville, Maryland  
OROFACIAL MYOFUNCTIONAL DISORDERS IN CHILDREN

Richard K. Peach  
Rush University and Rush-Presbyterian-St. Luke’s Medical Center, Chicago, Illinois  
APHASIA, GLOBAL  
MELODIC INTONATION THERAPY

Adrienne L. Perlman  
University of Illinois at Urbana-Champaign, Champaign, Illinois  
DYSPHAGIA, ORAL AND PHARYNGEAL

John M. Pettit  
Radford University, Radford, Virginia  
TRANSEXUALISM AND SEX REASSIGNMENT: SPEECH DIFFERENCES

James O. Pickles  
The University of Queensland, Brisbane, Australia  
PHYSIOLOGICAL BASES OF HEARING

Geoff Plant  
The Hearing Rehabilitation Foundation, Somerville, Massachusetts  
SPEECH TRACKING

Sheila Pratt  
University of Pittsburgh, Pittsburgh, Pennsylvania  
AUDITORY TRAINING  
SPEECH DISORDERS SECONDARY TO HEARING IMPAIRMENT ACQUIRED IN Adulthood

Adele Proctor  
University of Illinois at Urbana-Champaign, Champaign, Illinois  
DIALECT, REGIONAL

Jennie Pyers  
University of California at Berkeley, Berkeley, California  
LANGUAGE OF THE DEAF: SIGN LANGUAGE

Lorraine Olson Ramig  
University of Colorado, Boulder, Colorado  
HYPOKINETIC LARYNGEAL MOVEMENT DISORDERS

Steven Z. Rapcsak  
University of Arizona, Tucson, Arizona  
AGRAPHIA

Sean M. Redmond  
University of Utah, Salt Lake City, Utah  
SOCIAL DEVELOPMENT AND LANGUAGE IMPAIRMENT

Charlotte M. Reed  
Massachusetts Institute of Technology, Cambridge, Massachusetts  
FREQUENCY COMPRESSION

Joanne E. Roberts  
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina  
OTITIS MEDIA: EFFECTS ON CHILDREN’S LANGUAGE

Margaret A. Rogers  
University of Washington, Seattle, Washington  
APHASIA, PRIMARY PROGRESSIVE

Mary Ann Romski  
Georgia State University, Atlanta, Georgia  
AUGMENTATIVE AND ALTERNATIVE COMMUNICATION: GENERAL ISSUES  
MENTAL RETARDATION

Clark A. Rosen  
University of Pittsburgh, Pittsburgh, Pennsylvania  
ASSESSMENT OF FUNCTIONAL IMPACT OF VOICE DISORDERS

John C. Rosenbek  
College of Health Professions, Gainesville, Florida  
MUTISM, NEUROGENIC

Jackson Roush  
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina  
HEARING LOSS SCREENING: THE SCHOOL-AGE CHILD

Nelson Roy  
University of Utah, Salt Lake City, Utah  
FUNCTIONAL VOICE DISORDERS  
PSYCHOGENIC VOICE DISORDERS: DIRECT THERAPY

Dennis M. Ruscello  
West Virginia University, Morgantown, West Virginia  
PHONOLOGICAL ERRORS, RESIDUAL

Ron Scherer  
Bowling Green State University, Bowling Green, Ohio  
VOICE PRODUCTION: PHYSICS AND PHYSIOLOGY

Robert S. Schlauch  
University of Minnesota, Minneapolis, Minnesota  
PSEUDOHYPOPACUSIS
Carson T. Schütze
University of California at Los Angeles, Los Angeles, California
Morphosyntax and Syntax

Rose A. Sevcik
Georgia State University, Atlanta, Georgia
Augmentative and Alternative Communication: General Issues
Mental Retardation

Harry N. Seymour
University of Massachusetts at Amherst, Amherst, Massachusetts
Language Disorders in African-American Children

Janet E. Shanks
Department of Veterans Affairs Medical Center, Long Beach, California; University of California at Irvine, Irvine, California
Tympanometry

Robert V. Shannon
House Ear Institute, Los Angeles, California
Auditory Brainstem Implant

Lewis P. Shapiro
San Diego State University, San Diego, California
Argument Structure: Representation and Processing

Elaine R. Silliman
University of Florida, Tampa, Florida
Inclusion Models for Children with Developmental Disabilities

Yvonne S. Sininger
House Ear Institute, Los Angeles, California
Auditory Neuropathy in Children

Bernadette SKa
University of Montreal, Montreal, Quebec
Discourse Impairments

Joseph J. Smaldino
University of Northern Iowa, Cedar Falls, Iowa
Classroom Acoustics

Steven L. Small
University of Chicago, Chicago, Illinois
Aphasia Treatment: Pharmacological Approaches

Ann Bosma Smit
Kansas State University, Manhattan, Kansas
Speech Sampling, Articulation Tests, and Intelligibility in Children with Phonological Errors
Speech Sampling, Articulation Tests, and Intelligibility in Children with Residual Errors

Pamela E. Souza
University of Washington, Seattle, Washington
Suprathreshold Speech Recognition

Charles Speaks
University of Minnesota, Minneapolis, Minnesota
Dichotic Listening

Joy Stockhouse
University of Sheffield, Sheffield, United Kingdom
Speech Disorders in Children: A Psycholinguistic Perspective

Elaine T. Stathopoulos
SUNY-Buffalo, Buffalo, New York
Voice Disorders in Children

Joseph Stemple
Institute for Voice Analysis and Rehabilitation, Dayton, Ohio
Voice Therapy: Holistic Techniques

Kenneth N. Stevens
Massachusetts Institute of Technology, Cambridge, Massachusetts
Voice Acoustics

Ida J. Stockman
Michigan State University, East Lansing, Michigan
Phonology: Clinical Issues in Serving Speakers of African-American Vernacular English

Carol Stoel-Gammon
University of Washington, Seattle, Washington
Mental Retardation and Speech in Children

Edythe A. Strand
Mayo Clinic, Rochester, Minnesota
Dysarthrias: Management

Kathy Stratman
Wichita State University, Wichita, Kansas
Phonological Awareness Intervention for Children with Expressive Phonological Impairments

Gerald A. Studebaker
University of Memphis, Memphis, Tennessee
Speech Perception Indices

Johan Sundberg
KTH Voice Research Centre, Stockholm, Sweden
The Singing Voice

Anne Marie Tharpe
Vanderbilt University Medical Center, Nashville, Tennessee
Audition in Children, Development of Hearing Loss and Teratogenic Drugs or Chemicals

Jack E. Thomas
Mayo Clinic, Rochester, Minnesota
Laryngectomy

Connie A. Tompkins
University of Pittsburgh, Pittsburgh, Pennsylvania
Right Hemisphere Language Disorders

Robert G. Turner
Louisiana State University Health Sciences Center, New Orleans, Louisiana
Clinical Decision Analysis

Ann A. Tyler
University of Nevada at Reno, Reno, Nevada
Speech Disorders in Children: Speech-Language Approaches

Richard S. Tyler
University of Iowa, Iowa City, Iowa
Cochlear Implants in Adults: Candidacy

Hanna K. Ulatowska
University of Texas at Dallas, Dallas, Texas
Discourse

Miodrag Velickovic
The Mount Sinai Medical Center, New York
Hyckokinetic Laryngeal Movement Disorders

Shelley L. Velleman
University of Massachusetts at Amherst, Amherst, Massachusetts
Speech Disorders in Children: Descriptive Linguistic Approaches
Katherine Verdolini
University of Pittsburgh, Pittsburgh, Pennsylvania
VOICE THERAPY FOR ADULTS
VOICE THERAPY FOR PROFESSIONAL VOICE USERS

Steven F. Warren
University of Kansas, Lawrence, Kansas
PRELINGUISTIC COMMUNICATION INTERVENTION FOR CHILDREN WITH DEVELOPMENTAL DISABILITIES

Ruth V. Watkins
University of Illinois at Urbana-Champaign, Champaign, Illinois
LANGUAGE IN CHILDREN WHO STUTTER

Peter Watson
Case Western Reserve University, Cleveland, Ohio
VOICE THERAPY: BREATHING EXERCISES

Susan Ellis Weismer
University of Wisconsin-Madison, Madison, Wisconsin
MEMORY AND PROCESSING CAPACITY

Nathan V. Welham
University of Wisconsin-Madison, Madison, Wisconsin
INSTRUMENTAL ASSESSMENT OF CHILDREN’S VOICE

Robert T. Wertz
VA Tennessee Valley Healthcare System, Nashville, Tennessee
APHASIA: THE CLASSICAL SYNDROMES

Susan L. Whitney
University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania
VESTIBULAR REHABILITATION

Judith E. Widen
University of Kansas Medical Center
OTOACOUSTIC EMISSIONS IN CHILDREN

Terry L. Wiley
Arizona State University, Tempe, Arizona
PURE-TONE THRESHOLD ASSESSMENT

Jennifer Windsor
University of Minnesota, Minneapolis, Minnesota
LANGUAGE DISORDERS IN SCHOOL-AGE CHILDREN: OVERVIEW

Shelley Witt
University of Iowa, Iowa City, Iowa
COCHLEAR IMPLANTS IN ADULTS: CANDIDACY

Diane M. Wrisley
Oregon Health and Sciences University, Beaverton, Oregon
VESTIBULAR REHABILITATION

Michael K. Wynne
Indiana University School of Medicine, Indianapolis, Indiana
PEDIATRIC AUDIOLOGY: THE TEST BATTERY APPROACH

William S. Yacullo
Governors State University, University Park, Illinois
 MASKING

J. Scott Yaruss
University of Pittsburgh, Pittsburgh, Pennsylvania
SPEECH DISFLUENCY AND STUTTERING IN CHILDREN

Mehmet Yavas
Florida International University, Miami, Florida
BILINGUALISM, SPEECH ISSUES IN

Paul J. Yoder
Vanderbilt University, Nashville, Tennessee
PRELINGUISTIC COMMUNICATION INTERVENTION FOR CHILDREN WITH DEVELOPMENTAL DISABILITIES

Christine Yoshinaga-Itano
University of Colorado, Boulder, Colorado
ASSESSMENT OF AND INTERVENTION WITH CHILDREN WHO ARE DEAF OR HARD OF HEARING

William A. Yost
Loyola University Chicago, Chicago, Illinois
AUDITORY SCENE ANALYSIS
PITCH PERCEPTION
Dillon, Harvey, 419, 420, 481, 482–485
Dillon, J. M., 240
DiMattia, M. S., 89
Ding, D. L., 509
Dinnsen, D. A., 345
Dinnunen, D. A., 174, 199
Dionne, G., 371
Dirks, D. D., 501, 502, 549
Ditelsberg, J. S., 56
Divenyi, P. L., 109
Dix, R., 467
Dixon, R., 475
Dixon-Ernst, C., 499
Dobie, R. A., 508, 509, 510, 557
Doeherty, G., 49, 50
Doctor, R. F., 224
Dodd, B., 120, 122, 142, 143, 189, 190, 191, 218–219, 336
Dodd, B. J., 141, 281
Dodd-Murphy, J., 495
Dodds, B., 43
Dodds, L., 417
Dodds, L. W., 509
Döerfler, L., 476
Doevshuk, C., 176, 179
Doggett, D., 133
Dogil, G., 102
Doherty, T., 92
Dolan, T., 424
Dollaghan, Christine A., 135–136, 350, 359, 371, 403
Dolly, J. O., 38, 39
Dommeguies, J.-Y., 394
Don, M., 433
Donahue-Kilburg, G., 285, 286
Donaldson, G. S., 551
Donlan, C., 184, 185, 350, 400, 402
Donnellan, A., 116
Doolan, S. J., 157
Döpke, S., 280
Dorkos, V. P., 472
Dorn, P. A., 513, 515
Doupe, A. J., 275
Dove, H., 56, 86, 90, 96
Dow, R. S., 315
Downhill, J. R., Jr., 258
Downs, J. A., 47
Dowty, D., 270
Doyle, M., 111
Doyle, Patrick J., 101–103, 104–106
Doyle, Philip C., 10–12, 45–48, 80–81
Doyle, T. N., 532
Drachman, D., 240
Drager, K., 113
Drake, A., 476
Draper, M., 7
Dravins, L. W., 509
Driscoll, C., 31, 39
Dromi, E., 332
Dronkers, Nina F., 60, 102, 184, 185, 350, 400, 402
Drouin-Casson, G., 291
Drury, L. K., 371
Dukert, L. G., 518
Duckworth, M., 150
Duchaux, R., 512
Dudley, H., 473
Duffy, R. J., 303
Dukette, D., 400
Dulon, D., 519
Dunca, P. V., 258
Duncan, G. W., 101, 252, 264
Duncan, P. W., 564
Dune, P. E., 226
Dunn, C., 198, 213
Dunn, D. D., 508
Dunn, J. W., 505
Dunwoody, G. W., 146
Duong, Anh, 302–304
Durieux-Smith, A., 430
Durkin, L., 39
Durlach, N. I., 419, 472, 473, 541
Durand, V. M., 573
Duranta, J., 430, 462
Durst, E., 102
Duyao, M. P., 427
Dworkin, J. P., 45, 48, 105, 202
Dye, L. M., 455
Dyer, J. J., 493, 494
Dykens, E., 310, 403
Dykstra, K., 197
Dzul, A., 34
E
Eadie, Tanya L., 10–12
Eagle, E., 495–496
Eales, C., 261
Earle, A. S., 21
Easton, D. L., 206
Ebendal, T., 258
Eckert, R. C., 38
Eddington, D. K., 449
Eddins, David A., 553–555
Eddy, W. F., 272, 311, 313
Edelman, S. W., 309
Eden, G. F., 329
Edgerton, B., 208
Edgerton, B. W., 427
Edgerton, M. T., 224
Edmundson, A., 143
Edwards, B., 481
Edwards, Brent, 420
Edwards, J., 184, 350
Edwards, Mary Louise, 150–152, 199, 205, 213
Edwards, S., 141
Edwards, V. T., 329, 402
Egger, J., 29
Ehrlich, J. S., 303
Eichen, E., 342
Eigsti, I.-M., 345
Eilers, R. E., 151, 370, 425
Eimas, P., 425
Eisele, J. A., 312, 396
Eisenberg, H. M., 396
Eisenberg, L. S., 427, 455
Eisner, M. P., 137
Ekelman, B. L., 184, 312
Elbhart, H., 283
Elbaum, B., 309
Elberling, C., 488
Elbert, M., 198, 199, 205, 214
Eldert, E., 548
Eldredge, L., 494
Newby, H., 475
Newman, Arlene, 419
Newman, C. W., 21, 86, 470, 557
Newman, L., 258
Newman, J. D., 59, 60
Newman, R., 395
Newport, E., 341, 342
Newton, R. A., 564
Ng. J., 55
Niccacci, N., 458, 459, 460
Ni Chasaide, A., 6
Nichol, H., 29, 50
Nicholas, J., 224
Oates, J. M., 224
Oates, J., 95
Nuutila, A., 368
Nuttal, A. L., 14
Nuss, R. C., 72, 73
Nuss, D., 72, 73
Nuss, D., 72, 73
Nuttal, A. L., 14
Nuutila, A., 368
Oates, J., 95
Nuss, R. C., 72, 73
Nuss, D., 72, 73
Nuss, R. C., 72, 73
Nuttal, A. L., 14
Nuutila, A., 368
Subject Index

A

AAA (American Academy of Audiology), on screening for hearing loss, 495
AAC. See Augmentative and alternative communication (AAC)
Aachener Aphasie Test (AAT), 253
AAE. See African-American English (AAE)
AAVE (African American Vernacular English). See African-American English (AAE)
Abdomen, in speech production, 83
Abdominal pneumobels, speech production with, 226
Adductor spasmodic dysphonia (ABSD), botulinum toxin for, 38
ABI (auditory brainstem implant), 427–429
ABI. See Auditory Brainstem Response (ABR)
Absolute gradient, of tympanogram, 506
Absolute gradient, of tympanogram, 506
Absolute pitch, 145
Absolute prosody, 108, 381
African-American children, language disorders in, 318–320
African-American English (AAE) camouflaged forms in, 296
African-American Vernacular English (AAVE). See African-American English (AAE)
Adolescence, language disorders in, 327
Adoption studies, of speech disorders, 185
AD (Adductor spasmodic dysphonia), botulinum toxin for, 38–39
Adults auditory brainstem response in, 429–433
communication disorders in, 283–284
voice therapy for, 88–90
Aerodynamic assessment of children’s voice, 36–37, 67–71
instrumentation for, 170
of vocal function, 7–9
Aerometry, in speech assessment, 170
Affect, flat, 108
Affective prosody, 108, 381
African-American children, language disorders in, 318–320
African-American English (AAE) camouflaged forms in, 296
contrastive patterns in, 295, 298–299
language acquisition milestones for, 318–319
phonologic issues with, 158–160
African Vernacular English (AAVE). See African-American English (AAE)
Aging hearing dysfunction related to, 527–530
masking-level difference with, 490
noise-induced hearing loss and, 510
voice disorders of, 72–74
voice therapy for, 91–94
Agrammatism, 231–232
comparative aphasia of, 265–268
“least pronunciation effort” approach to, 265–266
reversibility/mapping disorders in, 383–384
trace deletion hypothesis of, 231, 271, 407–409
tree pruning hypothesis of, 405–407
Agraphia, 233–235
Air, pressure
in children, 36–37, 67–70
Airflow alternating glottal, 69
average oral, 68
measurement of, 7–8
in children, 36–37, 67–70
Airflow open quotient, 68
in children, 69, 70
Air pressure intraoral, 68
measurement of, 8, 9
in children, 36–37, 68
subglottal, 68
Air volumes, measurement of, 7
Airway resistance, laryngeal, 68
in children, 70–71
and sound pressure level, 70–71
Akinetic mutism (AM), 145
Alaryngeal voice, 10–12, 138–139
Alcohol abuse, during pregnancy, 194–195
hearing loss due to, 494
Alexia, 226–239
Aliasing, 472
Allergies, laryngitis due to, 34
Alcoholic conversion, 234
Allographic conversion, 235
Allophonic variation, 174
Alouatta spp, vocal production system in, 57, 58
Alport syndrome, hearing loss in, 479
ALS (amyotrophic lateral sclerosis) augmentative and alternative communication approaches in, 111
dysarthria due to, 126
Alström disease, hearing loss in, 479
Alternating glottal airflow, 69
in children, 69–70
sound pressure level and, 69–70
Alternating motion rates (AMRs), in dysarthria, 126
Alveolar defects of, 46
Alzheimer’s disease (AD) agrammatism in, 231, 240
communication disorders in, 291–292
language impairments in, 240–241
AM (akinetic mutism), 145
American Academy of Audiology (AAA), on screening for hearing loss, 495
American National Standards Institute (ANSI) on articulation index, 539
on calibration of instruments, 496
on hearing protectors, 498, 499
on noise-induced hearing loss, 508
American Sign Language (ASL), 339–343, 423
American Speech-Language-Hearing Association (ASHA) on audiometric symbols, 502
on dialects, 297
FACs of, 283
on inclusion, 308
on language disorders in African-American children, 318
on screening for hearing loss, 495, 561
Aminoglycosides, ototoxicity of, 493, 518
Amplitude compression, in hearing aids, 413–420
Aminoglycosides, ototoxicity of, 493, 518
and loudness, 413–416
and masking, 416–417
Amplitude modulation detection, 553, 554–555
Physiology of, 417–418
and sound quality, 488
Amplicon detection, 553, 554–555
AMRs (alternating motion rates), in dysarthria, 126
Amyloidosis, laryngeal involvement in, 34
Amyotrophic lateral sclerosis (ALS) augmentative and alternative communication approaches in, 111
dysarthria due to, 126
AN (auditory neuropathy), in children, 433–436
Analgesics, ototoxicity of, 519
Anarthric mutism, 145–146
Anomia, in primary progressive aphasia, 247, 248
Anomia, 240
Anosognosia, in Wernicke’s aphasia, 252
language acquisition with
for English, 336–338
sign, 339–343
Michel, 479
Montrul, 479
pure word, 252, 263
Scheibe, 479
Decision matrix, 444
Declarative, 375
Deep agraphia, 234, 235
Deep alexia, 238, 239
Deep dyslexia, 238, 239
Deer, vocal production system in, 58
Deficits of linguistic competence, 273
Dehydration, vocal hygiene for, 55, 89
Deictic terms, in autism, 116
Dehydration, vocal hygiene for, 55, 89
Dehydration, vocal hygiene for, 55, 89
Dehydration, vocal hygiene for, 55, 89
Dehydration, vocal hygiene for, 55, 89
Deep acquired, 238
attentional, 237
directive, 238, 239
neglect, 237
phonological, 238
surface, 238
Deep dyslexia, 238, 239
Dysphonia(s)
Dysphasia, oral and pharyngeal, 132–134
Dysphonia(s)
functional
direct therapy for, 49–51
etiology of, 27–29
muscular tension
botulinum toxin for, 38
voice handicap assessment in, 22
paradoxical breathing dysphonia, botulinum
toxin for, 38
spasmodic
acoustic assessment of, 4–5
botulinum toxin for, 38–39
laryngeal reinnervation for, 43
Dyspraxia. See Developmental apraxia of
speech (DAS)

E
Ear, physiology of, 522–525
Ear advantage, 458–460
Ear canal volume, 505
Eardrops, ototoxicity of, 518–519
Ears, 498, 499
Earplugs, 497–498, 499
Earphones, 500–501, 534, 535
Earplugs, 497–498, 499
E-A-TONE 3A earphones, 501
Ebonics. See African-American English
(AAE)
Echosalia, in autism, 116
ECochG (electrocochleography), 461–465
Edema, Reinke’s, and voice disorders in
elderly, 72–73
EEG (electroencephalography), 305
Efficiency (EF), 447
EGG (electroglottography), 23–26
of children’s voice, 27
Eight-speaker localization test, 453
Elderly
hearing dysfunction in, 527–530
masking-level difference with, 490
noise-induced hearing loss and, 510
voice disorders of, 72–74
voice therapy for, 91–94
Electrical stimulation, for tinnitus, 557
Electrocochleography (ECochG), 461–465
Electrode(s)
in cochlear implant, 448
 tympanic membrane, 462
Electrocochleography (EEG), 305
Electroglottography (EGG), 23–26
of children’s voice, 27
Electrolaryngography, 23–26
Electrolarynx, 12, 138
Electromagnetic articulography (EMA), 171
Peripheral laryngeal nerve damage, in elderly, 72
Peripheral neuropathy, with auditory neuropathy, 434
Peripheral paralyis, in elderly, 72
Peripheral structural changes due to surgical ablation, 45–48
due to trauma, 45
Perisylvian cortex, in vocalization, 60, 61
Perseveration, 361–366
Pervasive developmental disorders (PDDs), 116
augmentative and alternative communication approaches with, 114
semantic development with, 395–397
PET (positron emission tomography), 305–306
of stuttering, 222
of subcortical involvement, 317
PET (pressure equalization tube), tympanometry of, 561
Pharmacological approaches, to aphasia, 25
Pharyngoepithelial (PEO) segment, 10–11
Pharynx, of infants and young children, 177
Phoneme-grapheme conversion, 234, 235
Phonemic versions, 413
Phonation. See also Voice evolution of, 56–58
flow, 52
hyperfucntional, 52
hypofunctional, 52
Phonemic carryover perseveration, 362
Phonemic contrast, with hearing loss, 208
Phonemic development with mental retardation, 141
Phonemic inventory, 275, 363–365
Phonetic inventory, 104
Phonetic organization, 143
Phonetic place, integral stimulation and, 104
Phonetic transcription, of speech in children, 150–152
Phonological acquisition, delayed, 219
Phonological agraphia, 234–235
Phonological alexia, 238
Phonological analysis, of language disorders in aphasia, 363–365
Phonological awareness, in children, 190, 324
Phonological awareness intervention, for children with expressive phonological impairments, 153–155
Phonological breakdown, in aphasia, 363, 364
Phonological development, of Latino children, 210–212
Phonological disorders
of articulation disorders, 196
in bilingual Latino children, 211–212
description and classification of, 219
developmental, 156
Pharmacological approaches, to aphasia, 25
Pharmacological errors in aphasia, 366–368
children with, speech sampling, articulation tests, and intelligibility in, 213–214
residual, 156–157
Phonological manipulation, 155
Phonological process analysis, 214
Phonological processing, 153
and reading disability, 329
Phonological process therapy, 199
Phonological sensitivity, 153
Phonologic issues, with speakers of African-American Vernacular English, 158–160
Phonologic lexicon in aphasia, 234
in alexia, 236, 237–238
Phonology in bilingualism, 279
linear and nonlinear, 175, 199
Phonology testing, of children with phonological errors, 214
with residual errors, 217
Photo Articulation Test—Third Edition, 216
Phrasal metaphor, right hemisphere in, 387
Phrasal repetitions, in children, 180
Phrenic nerve, laryngeal reinnervation with, 43
Phrenic nerve paresis, speech production with, 226
Pierre Robin syndrome, hearing loss in, 479
Pinna, physiology of, 522
Piracetam, for aphasia, 258
Pitch
in aprosodias, 109
in auditory development, 425
complex, 526–527
factors leading to, 77
of missing fundamental, 438
musical, 52–53, 525–526
perception of, 525–527
Pitch chroma, 526
Pitch control, 75–76
Pitch height, 526
Pitch range, 53–54
Pitch scales, 526
PIW (Percentage of Intelligible Words), 214, 217
Place coding, of frequency, 523
Plasticity, constrained, 311
Plateau, 502
Playback, slow, 472
Phlebitis, in speech assessment, 170
PMT (prelinguistic milieu teaching), 376–377
POGO II procedure, for prescriptive fitting of hearing aid, 485
Point vowels, 59
Poliomyelitis, relapsing, laryngeal involvement in, 34
Polyphonic degeneration, and voice disorders in elderly, 72–73
Positional nystagmus, 469
Positive pressure ventilation in infants and young children, 177
speech production with, 226, 227
Positive reinforcement, 192
Positron emission tomography (PET), 305–306
of stuttering, 222
of subcortical involvement, 317
Vocal production system
breathing in, 82–84
evolution of, 56–58
physics and physiology of, 75–77
Vocal tract filter function, 76
Vocal tremor, 94
botulinum toxin for, 38–39
Vocoding, 473
Voice, 1–98
acoustic assessment of, 3–6
aryalance, 10–12, 138–139
in children, instrumental assessment of, 35–37
confidential, 93
electroglottographic assessment of, 23–26
esophageal, 138
resonant, 52, 89, 93
in singing, 51–54
Voice acoustics, 63–67
Voice Activity and Participation Profile (VAPP), 22
Voice breaks
intermittent, botulinum toxin for, 38
in voice exercises, 86–87
Voice conservation, for professional voice users, 96
Voice disorders
of aging, 72–74
assessment of functional impact of, 20–23
in children, instrumental assessment of, 35–37, 67–71
functional (psychogenic)
direct therapy for, 49–51
etiology of, 27–29
in parkinsonism, 30–31
prevalence of, 20
psychogenic, 27–29
Voice Handicap Index (VHI), 21–22
Voice handicap measures, 20–23
Voice Outcome Survey (VOS), 22
Voice production system
evolution of, 56–58
physics and physiology of, 75–77
Voice quality
after conservation laryngectomy, 81
factors leading to, 77
perceptual evaluation of, 78–79
Voice rehabilitation, after laryngectomy, 10–12
partial or conservation, 80–81
Voice source, during singing, 51, 52–53
Voice therapy
for adults, 88–90
breathing exercises in, 82–84
for functional dysphonia, 49–51
holistic, 85–87
for neurological aging-related voice disorders, 91–94
for professional voice users, 95–97
for transsexuals, 224, 225
Volume velocity, 63, 64
Volume-velocity waveform, 63, 65
VOR (vestibulo-ocular reflex), 565
VOS (Voice Outcome Survey), 22
Vowel(s), point, 58
Vowel(s), production, 143
Vowel prolongation, in dysarthria, 126
Whole-word repetitions, in children, 180
Whole-word carryover, 362
Whole-word repetitions, in children, 180
Whole-band spectrogram, 171–173
Wide dynamic range compression (WDRC), 413
Williams syndrome, and speech, 141
Wilson Voice Profile System, 78
Word deafness, pure, 252, 263
Word use, in autism, 116
Working memory
in school-age children, 328
and semantic development, 395–396
in specific language impairment, 349–351
Writing disorders, 233–235
X
Xerostomia, due to radiotherapy, 46
X-ray microbeam imaging, in speech assessment, 171
Y
Yawn-sign phonation, 88, 89
Yes/no target-monitoring task, 458–459
Z
Zero-crossing-rate division, 473
Waardenburg syndrome, hearing loss in, 479
WAB (Western Aphasia Battery), 253