

SPRINGER HANDBOOK OF AUDITORY RESEARCH

Series Editors: Richard R. Fay and Arthur N. Popper

Lynne Werner
Richard R. Fay
Arthur N. Popper
Editors

Human Auditory Development

 Springer

Springer Handbook of Auditory Research

For further volumes:
<http://www.springer.com/series/2506>

Lynne A. Werner • Richard R. Fay
Arthur N. Popper
Editors

Human Auditory Development

With 47 Illustrations

 Springer

Editors

Lynne A. Werner
Department of Speech and Hearing Sciences
University of Washington
Seattle, WA 98105-6246, USA
lawerner@u.washington.edu

Richard R. Fay
Marine Biological Laboratory
Woods Hole, MA 02543, USA
rfay@luc.edu

Arthur N. Popper
Department of Biology
University of Maryland
College Park, MD 20742, USA
apopper@umd.edu

ISSN 0947-2657

ISBN 978-1-4614-1420-9

e-ISBN 978-1-4614-1421-6

DOI 10.1007/978-1-4614-1421-6

Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011941593

© Springer Science+Business Media, LLC 2012

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)



This volume is dedicated to our friend and colleague Edwin W Rubel in appreciation of his science, his mentorship, and his support of research in all aspects of auditory development

Series Preface

The Springer Handbook of Auditory Research presents a series of comprehensive and synthetic reviews of the fundamental topics in modern auditory research. The volumes are aimed at all individuals with interests in hearing research including advanced graduate students, post-doctoral researchers, and clinical investigators. The volumes are intended to introduce new investigators to important aspects of hearing science and to help established investigators to better understand the fundamental theories and data in fields of hearing that they may not normally follow closely.

Each volume presents a particular topic comprehensively, and each serves as a synthetic overview and guide to the literature. As such, the chapters present neither exhaustive data reviews nor original research that has not yet appeared in peer-reviewed journals. The volumes focus on topics that have developed a solid data and conceptual foundation rather than on those for which a literature is only beginning to develop. New research areas will be covered on a timely basis in the series as they begin to mature.

Each volume in the series consists of a few substantial chapters on a particular topic. In some cases, the topics will be ones of traditional interest for which there is a substantial body of data and theory, such as auditory neuroanatomy (Vol. 1) and neurophysiology (Vol. 2). Other volumes in the series deal with topics that have begun to mature more recently, such as development, plasticity, and computational models of neural processing. In many cases, the series editors are joined by a co-editor having special expertise in the topic of the volume.

Richard R. Fay, Falmouth, MA
Arthur N. Popper, College Park, MD

Volume Preface

This volume deals with what is currently known about the development of hearing and the auditory system in humans. The chapters in the volume are arranged roughly in “bottom-up” order, beginning with the development of the ear, moving on to the development of the auditory nervous system, and then to the development of various aspects of perception and cognition that depend on the development of the ear and auditory nervous system.

In Chapter 2, Abdala and Keefe summarize what is known about the postnatal development of the ear, including both the conductive apparatus and the cochlea. They discuss in detail the relationship between conductive maturation and the response of the cochlea as recorded in the ear canal as otoacoustic emissions. In Chapter 3, Eggermont and Moore consider the morphological and physiological development of the human auditory system, relating landmarks in the development of auditory evoked potentials to structural changes in the auditory pathways in infancy, childhood, and adolescence. Buss, Hall, and Grose (Chapter 4) then discuss the development of basic auditory capacities—sensitivity, spectral resolution, intensity resolution, and temporal resolution—with a view toward disentangling the contributions of sensory and nonsensory factors. Leibold (Chapter 5) picks up where Chapter 4 left off, and discusses how children deal with complex listening situations, focusing on auditory scene analysis and auditory selective attention. Litovsky, in Chapter 6, describes the development of spatial hearing, a capacity that depends on the status of basic auditory capacities, as well as specialized central processing, and that in turn subserves auditory scene analysis as well as selective attention.

Chapters 7 and 8 address the endpoint of audition, the extraction of information or meaning from the neurally encoded, analyzed, and selected auditory stream. The ability to process the information contained in speech is dealt with by Panneton and Newman (Chapter 7), while the ability to process information contained in music is discussed by Trainor and Urau (Chapter 8). Finally, Eisenberg, Johnson, Ambrose, and Martinez (Chapter 9) consider the effects of auditory deprivation or degraded auditory input on the development of speech perception in deaf and hard-of-hearing children.

As with all volumes in the Springer Handbook of Auditory Research series, there is often related material in other volumes that inform material in the current book. In particular, this volume holds many parallels to issues raised about hearing in older humans in *The Aging Auditory System* (Vol. 34, edited by Gordon-Salant, Frisina, Popper, and Fay). General development of the auditory system is discussed in many chapters in *Development of the Ear* (Vol. 25, edited by Kelley, Wu, Popper, and Fay) and in *Plasticity of the Auditory System* (Vol. 22, edited by Parks, Rubel, Popper, and Fay). Issues related to general perception of sound by humans are considered at length in *Auditory Perception of Sound Sources* (Vol. 29, edited by Yost, Popper, and Fay) and in an earlier volume in the series on *Human Psychophysics* (Vol. 3, also edited by Yost, Popper, and Fay). Chapters in the current volume on music perception are expanded upon in *Music Perception* (Vol. 36, edited by Jones, Fay, and Popper) and on sound localization in *Sound Source Localization* (Vol. 25, edited by Popper and Fay). Finally, issues related to hearing loss and use of cochlear implants are considered at length in two SHAR volumes, *Auditory Prostheses* (Vol. 18, edited by Zeng, Popper, and Fay) and in *Auditory Prostheses: New Horizons* (Vol. 39, also edited by Zeng et al.).

Lynne A. Werner, Seattle, WA
Richard R. Fay, Falmouth, MA
Arthur N. Popper, College Park, MD

Contents

1 Overview and Issues in Human Auditory Development.....	1
Lynne A. Werner	
2 Morphological and Functional Ear Development.....	19
Carolina Abdala and Douglas H. Keefe	
3 Morphological and Functional Development of the Auditory Nervous System	61
Jos J. Eggermont and Jean K. Moore	
4 Development of Auditory Coding as Reflected in Psychophysical Performance	107
Emily Buss, Joseph W. Hall III, and John H. Grose	
5 Development of Auditory Scene Analysis and Auditory Attention.....	137
Lori J. Leibold	
6 Development of Binaural and Spatial Hearing	163
Ruth Y. Litovsky	
7 Development of Speech Perception.....	197
Robin Panneton and Rochelle Newman	
8 Development of Pitch and Music Perception.....	223
Laurel J. Trainor and Andrea Unrau	
9 Atypical Auditory Development and Effects of Experience.....	255
Laurie S. Eisenberg, Karen C. Johnson, Sophie E. Ambrose, and Amy S. Martinez	
Index.....	279

Contributors

Carolina Abdala Division of Communication & Auditory Neuroscience,
House Research Institute, Los Angeles, CA, USA
cabdala@hei.org

Sophie E. Ambrose Children's Auditory Research & Evaluation Center,
House Research Institute, Los Angeles, CA, USA
sophie.ambrose@boystown.org

Emily Buss Department of Otolaryngology/Head and Neck Surgery,
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
ebuss@med.unc.edu

Jos J. Eggermont Department of Physiology and Pharmacology,
Department of Psychology, University of Calgary, Calgary, AB, Canada
eggermon@ucalgary.ca

Laurie S. Eisenberg Children's Auditory Research & Evaluation Center,
House Research Institute, Los Angeles, CA, USA
leisenberg@hei.org

John H. Grose Department of Otolaryngology/Head and Neck Surgery,
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
jhg@med.unc.edu

Joseph W. Hall III Department of Otolaryngology/Head and Neck Surgery,
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
jwh@med.unc.edu

Karen C. Johnson Children's Auditory Research & Evaluation Center,
House Research Institute, Los Angeles, CA, USA
kjohnson@hei.org

Douglas H. Keefe Boys Town National Research Hospital, Omaha, NE, USA
Douglas.Keefe@boystown.org

Lori J. Leibold Department of Allied Health Sciences, The University
of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, NC, USA
leibold@med.unc.edu

Ruth Y. Litovsky University of Wisconsin Waisman Center, Madison, WI, USA
Litovsky@waisman.wisc.edu

Amy S. Martinez Children's Auditory Research & Evaluation Center,
House Research Institute, Los Angeles, CA, USA
amartinez@hei.org

Jean K. Moore House Ear Institute, Los Angeles, CA, USA
drjkm@gmail.com

Rochelle Newman Department of Hearing & Speech Sciences,
University of Maryland, College Park, MD, USA
rnewman1@umd.edu

Robin Panneton Department of Psychology, Virginia Tech,
Blacksburg, VA, USA
panneton@vt.edu

Laurel J. Trainor Department of Psychology, Neuroscience & Behaviour,
McMaster University, Hamilton, ON, Canada
LJT@mcmaster.ca

Andrea Unrau Department of Psychology, Neuroscience & Behaviour,
McMaster University, Hamilton, ON, Canada
unrauaj@mcmaster.ca

Lynne A. Werner Department of Speech and Hearing Sciences,
University of Washington, Seattle, WA, USA
lawerner@uw.edu

Chapter 1

Overview and Issues in Human Auditory Development

Lynne A. Werner

1 Introduction

Broadly stated, the purpose of auditory processing is to reconstruct a scene of auditory objects, each of which can be identified and from each of which meaning can be extracted. The goal of auditory development, then, is to allow the mature system to accomplish that purpose accurately and efficiently. Auditory processing can be conceptualized as occurring in a series of stages, beginning with the encoding of sound at the ear and ending with recognition or understanding. Those who study auditory development describe whether and how the results of this final stage of processing improve with age, attempt to determine which stages of processing limit performance in immature listeners, and investigate the factors responsible for age-related change.

The sequence of events in this volume follows the sequence of events in sound processing: Sound is conducted to the inner ear, where it is encoded as a sequence of action potentials. This sequence in each auditory nerve fiber represents the waveform of sound in a third-octave frequency band. At low frequencies, both the temporal fine structure (TFS) and the envelope of the sound are represented, whereas at high frequencies, only the envelope is represented. The response properties of the cochlea are the primary determinants of the representation sent to the auditory brain, although the intrinsic response properties of primary auditory neurons are now understood to contribute (e.g., Adamson et al. 2002). The auditory nerve response represents the sum of all incoming sounds. From this representation, the auditory system extracts the features of sound that are critical to segregating the responses to different auditory objects and, subsequently, to identification, recognition, and

L.A. Werner (✉)
Department of Speech and Hearing Sciences, University of Washington,
Seattle, WA, USA
e-mail: lawerner@uw.edu

extracting meaning. Ultimately, the auditory scene is reconstructed as a spatial array of sound-producing objects. Meaning can then be extracted from each of these objects. Each event in this sequence is presumably associated with processing in specific neural structures in the auditory pathway and with complex interactions among those structures. These complex interactions include so-called top-down processes, by which subcortical and cortical structures influence the way that sound is encoded, the features that are extracted from the peripheral input, and so on.

The chapters in this volume describe what is currently known about the development of these processes in humans. The chapters are arranged roughly in “bottom-up” order, beginning with the development of the ear, moving on to the development of the auditory nervous system, then to the development of various aspects of perception that depend on the development of the ear and nervous system. In Chap. 2, Abdala and Keefe summarize what is known about the postnatal development of the ear, including both the conductive apparatus and the cochlea. One of the more interesting controversies in this field is discussed at length: Are age-related changes in cochlear response, as indicated by otoacoustic emissions, due to changes in the cochlea, or are they consequences of the maturation of the external and middle ear? Although no final conclusion is reached in the chapter, it is of some interest that early auditory behavior can be limited by immaturity of what may be thought of as the simplest of auditory processes, the conduction of sound into the ear. It should be noted, in any case, that the age-related changes in cochlear response that Abdala and Keefe describe involve an enhancement of cochlear tuning in some conditions, a phenomenon that cannot account for limited spectral resolution in early life.

In Chap. 3, Eggermont and Moore consider the morphological and physiological development of the human auditory system, relating landmarks in the development of auditory evoked potentials to structural changes in the auditory pathways in infancy, childhood, and adolescence. This material provides a new framework for understanding the development of auditory behavior, marking a real advance in a field in which unspecified “central” maturation has heretofore been cited as an explanation for nearly every age-related improvement in auditory performance.

Chaps. 4–8 describe the age-related improvements in auditory performance that result from the structural and functional changes in the auditory pathways described in Chaps. 2 and 3. Buss et al. (Chap. 4) begin with an account of the development of basic auditory capacities: sensitivity, spectral resolution, intensity resolution, and temporal resolution. In adults, these abilities are thought of as being set in the periphery. That is, the sensitivity and resolution established in the ear are carried through subsequent stages of neural processing. This idea is consistent with the fact that disorders of the peripheral auditory system result in perceptual deficits in these basic capacities. One important conclusion reached by Buss et al. is that immature neural processing can limit the expression of mature peripheral processing, not only because an immature ability to sustain attention can influence psychophysical measures of hearing, but because immature neural sensory processing can introduce noise that degrades performance, particularly in complex listening situations.

Leibold (Chap. 5) picks up where Chap. 4 left off, dealing in some detail with how children deal with complex listening situations. Here, the assumption is made that

the sensory pathway provides the child with sufficient information to support the reconstruction of the auditory scene and the extraction of information from a specified auditory object or event. Leibold makes the case that insufficient attention has been given to the development of auditory scene analysis and auditory selective attention and the potential impact of immaturity of these systems on performance in even basic auditory tasks. That is, a simple listening situation for an adult may well be a complex listening situation for an infant or child.

Chap. 6 deals with a special topic in basic auditory capacities: spatial hearing. Litovsky details the major events in the development of sound localization as well as the ability to use spatial information to improve auditory stream segregation and the extraction of information from one such stream. Again, an important message is that complex listening situations can present special problems for auditory processing by children.

Chaps. 7 and 8 get to the heart of the matter: Once incoming sound is encoded, the auditory scene is analyzed and one stream of information selected for processing, how does the ability to extract information from that stream change with age? The ability to process the information contained in speech is dealt with by Panneton and Newman (Chap. 7), while the ability to process information contained in music is considered by Trainor and Urau (Chap. 8). Both chapters conclude that even very young infants are sensitive to some of the information carried in these spectrotemporally complex sounds, but that mature processing is achieved only late in childhood or even in late adolescence. The idea that the auditory system is prepared to deal with a certain sort of message, with experience with sound elaborating on the basic process, is common to the literatures on the development of speech and music, respectively. It may even be that the sort of information extracted from speech and music—the spectrotemporal patterns carrying an emotion-laden message—is similar early in infancy. Nonetheless, the fundamental differences in the structural details that carry meaning in speech and in music necessitate the subsequent divergence of their developmental courses.

A common theme in many of the chapters in this volume is how experience with sound influences the development of perception. This question is structured, here and in the field in general, in two ways: First, what are the consequences of lack of exposure to sound on auditory system structure and function, and ultimately, on the ability to process sound? Second, how does experience with sound guide the development of sound processing? Eggermont and Moore (Chap. 3) lay the groundwork for both approaches, correlating development of auditory cortex with the infant's ability to process spoken language specifically and using studies of the development of auditory evoked potentials in deaf children who have received cochlear implants as a tool to understand when and where afferent input is critical to auditory neural development. Litovsky (Chap. 6) takes up the question of the effects of lack of sound input, the first question, from a different perspective, asking whether and when reintroduction of input can support the development of spatial hearing in the case of children with cochlea implants. Is there an age beyond which the introduction of binaural stimulation can no longer drive the development of functional circuits in support of sound localization and subsequently the ability to separate auditory

streams? Eisenberg et al. (Chap. 9) also consider the development of speech perception in deaf children who have received cochlear implants. Is there an age before which introduction of sound input can allow the perception of speech to proceed along its normal course? Eisenberg et al. also consider in detail the consequences of less severe hearing loss on the development of speech perception, an issue that is difficult to address given that the input, either before or after intervention, is not uniform across children and may not even be known with any degree of certainty.

Panneton and Newman (Chap. 7) and Trainor and Urau (Chap. 8), on the other hand, elaborate on the ways in which experience with sound guides the development of perception in typically developing children. In the case of speech, there is a wealth of detail concerning this process, starting with the acoustic bases of infant's preference for certain sorts of speech to emerging differences in the processing of native and nonnative, heard and unheard, speech sounds to the use of phonotactic information, allophonic variation, and existing vocabulary to segment ongoing speech into meaningful units. In the case of music, less detailed information is available in that most studies have focused on major distinctions, like sensitivity to musical-scale structure. On the other hand, the study of music perception provides an opportunity to examine the effects of variations in the quantity and quality of experience with sound, an effect that is typically not amenable to study in speech perception.

The remainder of this chapter summarizes some major findings in human auditory development discussed in detail in the remaining chapters, and in particular, emphasizes the interrelatedness of the material presented in the rest of the book. This extends from the relationship between peripheral and central responses to the relationship between the structural and physiological properties of the auditory pathway and auditory perception to the relationship between basic aspects of auditory perception and complex perceptual processes. Another important theme is how experience with sound influences auditory development at all levels of the system and for all types of perception.

2 Overview of Auditory Development

2.1 *Encoding and Feature Extraction*

As discussed by Buss et al. (Chap. 4), age-related improvement in every aspect of auditory encoding has been reported. Absolute sensitivity to high-frequency sound improves rapidly early in infancy; sensitivity to low-frequency sound improves more slowly and into the school years (e.g., Sinnott et al. 1983; Nozza and Wilson 1984; Olsho et al. 1988; Trehub et al. 1988; Tharpe and Ashmead 2001). Detection in noise also improves between infancy and school age (Schneider et al. 1989), although the final stages of development of spectral resolution appear to be complete by about 6 months of age (Olsho 1985; Schneider et al. 1990; Spetner and Olsho 1990). Discrimination of high-frequency tones improves in early infancy to nearly adult-like levels at 6 months of age; this pattern is consistent with the idea that high-frequency

discrimination depends on the place code for frequency (Moore 1973) and that spectral resolution, the precision of the place code, is mature by 6 months. Temporal resolution, the ability to follow the envelope of sound, is immature in infants, but appears to be adult-like by the end of the preschool period (Wightman et al. 1989; Werner et al. 1992; Hall and Grose 1994; Trehub et al. 1995). Although the development of temporal fine structure (TFS) processing has been less studied, it appears that by 5 years of age, children are able to use TFS to identify speech as well as adults are (Bertoncini et al. 2009). Consistent with that observation, discrimination of low-frequency pure tones, which in adults depends on the precision of TFS coding, is not mature until the school years (Maxon and Hochberg 1982; Olsho et al. 1987). Intensity coding is arguably not mature until the school years (see Buss et al. 2006; Buss et al., Chap. 4).

Thus, it appears that some aspects of sound encoding mature during infancy, while others require several years to approach adult status. One issue, of course, is that psychoacoustic measures of performance can be influenced by higher order processing such as sustained attention, motivation, and memory; Buss et al. (Chap. 4) and Leibold (Chap. 5) discuss this issue in detail. Another way to approach the question is to consider whether concomitant age-related changes in the function of the neural components of the auditory system can be identified.

The encoding of sound takes place in the ear. Electrophysiological studies provide one way of assessing the functional status of the ear. For example, wave I of the auditory brain stem response (ABR) is generated by the auditory nerve, and thus reflects the output of the cochlea. The sound pressure level at which ABR wave I can be detected is typically reported to be approximately 15 dB higher in newborn infants than in adults (e.g., Rotteveel et al. 1987; Sininger et al. 1997), suggesting some peripheral immaturity at term birth. However, the earliest studies of auditory morphological development in humans suggested that the inner ear is mature by full-term birth (Bredberg 1968; Pujol and Lavigne-Rebillard 1995). Interestingly, although a few studies continue to address this topic (e.g., Bibas et al. 2000, 2008; Ng 2000), no evidence countering that suggestion has been reported. In fact, Abdala and Keefe (Chap. 2) argue that if cochlear mechanisms are involved in apparent age effects on distortion product otoacoustic emissions (DPOAE), those mechanisms are more finely tuned in newborn infants than they are in adults.

The extent to which early responses are modified by conductive immaturity has been assessed only relatively recently. This may seem odd in that anyone who has ever looked at an infant can see that the external ear is smaller than an adult's. The landmark studies by Keefe and colleagues (1993; 1994) demonstrated that a newborn's conductive apparatus transmits less sound energy than an adult's, and the infant – adult difference is about the correct magnitude to account for the age-related differences in evoked potential thresholds. Further, small differences in conductive efficiency persist into the school years (Okabe et al. 1988) and likely contribute to the observed age differences in absolute sensitivity. Abdala and Keefe (Chap. 2) point out that conductive inefficiency is a viable explanation for infant – adult differences in DPOAE growth and tuning. As these authors also point out, however, the precise contribution of conductive immaturity will depend on how sounds are presented to the ear, whether in a sound field, through a headphone, or in a probe sealed in the ear canal.

It is unlikely, then, that sound coding by the ear can explain infants' and children's immature performance on tests of basic sound coding abilities. Eggermont and Moore (Chap. 3) offer a model of auditory neural development that can guide the search for explanations. Based on both anatomical and electrophysiological measures, they see auditory development progressing through two rough stages. In the first, beginning at birth and extending to about 6 months of age, the auditory nerve and brain stem pathways undergo rapid development. Synaptic efficiency and increased myelination result in increased neural conduction speed and improved information transfer. At term birth, afferent synaptic transmission from the cochlea for frequencies above approximately 6,000 Hz remains immature (Eggermont 1991, Eggermont et al., 1996), providing another potential contributor to immature sensitivity at high frequencies early in the postnatal period. Refinement of neural projection patterns in the brain stem also appears to be responsible for improvements in spectral resolution observed between 3 and 6 months of age (Folsom and Wynne 1987; Abdala and Slinger 1996) and for increased precision in the representation of speech (Bertoncini et al. 1988; Bertoncini 1993). Similarly, although the acoustic response of the head and external ear limit the information available to young infants about the spatial location of sound sources (e.g., Clifton et al. 1988; Ashmead et al. 1991), the maturation of these brain stem auditory pathways certainly contributes to early development of sound localization, described by Litovsky (Chap. 5). Moreover, the precedence effect, the tendency to assign the location of a sound source on the basis of the earliest arriving binaural information, emerges during this period (reviewed by Clifton 1992 and by Litovsky, Chap. 5).

During this period of development, however, the thalamocortical auditory pathway is quite immature (Yakovlev and Lecours 1967; Kinney et al. 1988; Moore and Guan 2001; Pujol et al. 2006). Although evoked potentials originating in auditory cortex can be recorded in infants in these early months (Pasman et al. 1991), it is likely that these responses reflect widespread activation of layer I auditory cortex by axons of brain stem reticular formation neurons that receive inputs from the primary auditory brain stem pathway and by intrinsic cortical neurons (Marin-Padilla and Marin-Padilla 1982; Moore and Guan 2001). Thus, early behavioral and evoked cortical responses to sound may well be mediated by a different sensory pathway than are mature responses. Interestingly, years ago developmental psychologists studying the early development of sound localization noted that the responses of newborn infants toward a sound source were qualitatively different from those of a 7-month-old and suggested that a switch from subcortical to cortical control mechanisms occurred around that age (e.g., Clifton et al. 1981; Clifton 1992).

Eggermont and Moore (Chap. 3) refer to this early period of auditory development as the "discrimination" stage, in reference to the fact that even young infants demonstrate relatively sophisticated responses to changes in complex sounds such as speech (Panneton and Newman, Chap. 7) and music (Trainor and Unrau, Chap. 8). Thus, whatever limitations immature brain stem circuits impose on the representation of sound available to young infants, the information that is available supports responses to sounds that are important for language development and for perceptual development generally. Whether infants and adults are using the same information about sounds to discriminate among them is a question discussed in a subsequent section of this chapter.

Eggermont and Moore (Chap. 3) suggest that at 6 months a second stage of auditory development gets underway. At this age, the inputs to auditory cortical layer I from the brain stem reticular formation are markedly reduced and other intrinsic neurons disappear. Concurrently, myelination of the acoustic radiation proceeds apace and thalamic axons begin to develop their typical pattern of termination in auditory cortical layer IV. Maturation of this circuitry continues for the next 4 or 5 years. However, cortical development is not complete at 5 years; rather, axonal density in auditory cortical layers II and III, the targets of many association and callosal axons, continues to increase to 12 years and beyond (Moore and Guan 2001). Some obligatory evoked potentials originating in auditory cortex (e.g., N1, Paetau et al. 1995; Gomes et al. 2001; Gilley et al. 2005) emerge during this late period of development and the latencies of cortical evoked potentials generally approach adult values (summarized by Eggermont and Moore, Chap. 3).

Eggermont and Moore (Chap. 3) refer to this as the “perceptual” stage, in reference to the increasing influence of the acoustic environment on children’s responses to sound. For example, it is in the second half of the first year of life that infants’ responses to speech become biased toward the native language (e.g., Werker and Tees 1984; Kuhl 1991; Panneton and Newman, Chap. 7).

It is also during this period of development that the remaining aspects of auditory encoding become adult-like. The ability to respond to sounds on the basis of the temporal properties, envelope and TFS, matures. Detection in noise, intensity resolution, and the precision of sound localization approach adult levels of performance. An important question remaining to be considered is the nature of these improvements in sound encoding. Given that the codes—for example, the phase-locked response to the acoustic waveform—are established and maintained at the periphery and in the brain stem pathways, is auditory development in the period between 6 months and 5 years properly defined as the development of encoding? It is possible that the codes for many acoustic properties are available to the child, but the child has not yet learned to use those codes efficiently. It is even possible that precision of auditory coding at the brain stem level improves as a result of top-down influences as the child discovers the utility of certain sorts of acoustic information (e.g., Anderson and Kraus 2010).

2.2 *Auditory Scene Analysis and Attention*

A 6-month-old, according to Eggermont and Moore, is at a turning point: Behavioral and electrophysiological responses to sound will no longer be dominated by the brain stem reticular formation inputs to cortex, but are coming under the control of the thalamocortical pathways. Six-month-olds resolve frequencies as well as adults do (e.g., Olsho 1985). They hear complex pitch when the fundamental component of a complex is missing (e.g., Clarkson and Clifton 1985; Trainor and Urau, Chap. 8). Their ability to judge the relative onset times of two tones is similar to adults’ (Jusczyk et al. 1980). They can identify the location of a sound source with an accuracy of approximately 15° (Ashmead et al. 1987; Morrongiello 1988).

The ability to extract these features of sound—frequency, harmonicity, temporal onset, spatial location—are some of the features that people use to form auditory objects. How well infants and children are actually able to reconstruct auditory scenes has not been established (Leibold, Chap. 5).

As Leibold (Chap. 5) discusses, Bregman (1990) posited two stages of auditory scene analysis, a primitive stage in which common properties of sound (or “cues”) in different frequency regions are identified and a schema-based stage in which stored representations of sound, if any, are imposed on results of initial analyses. Clearly, sound encoding and feature extraction impose constraints on a listener’s ability to perform this analysis. These may well be an issue for auditory scene analysis early in infancy (Buss et al., Chap. 4). Given an adequate representation of incoming sound, however, the initial “primitive” analysis must involve additional processing. Recent work suggests that this analysis is carried out, at least in part, at the level of the auditory brain stem (Pressnitzer et al. 2008). Because brain stem circuitry is still developing during infancy (Eggermont and Moore, Chap. 3), it might be predicted that this early stage of auditory scene analysis will demonstrate age-related improvements during infancy. These could be manifest as the emergence of the ability to use certain cues or a progressive decrease in the magnitude of cues needed to segregate the sound components from a common source. Of course, Bregman’s model of auditory scene analysis includes a major role for experience with sound in the formation of “schemas,” leading to the expectation of continued development as the child’s experience with sound builds. For example, common temporal modulations and harmonicity allow listeners to separate the components of one voice from those of another. However, prior knowledge of habitual prosody and fundamental frequency of a familiar voice can make this process more efficient (Newman and Evers 2007). The development of thalamocortical, intracortical, and efferent auditory pathways may all be involved, not only in the ability to store the schemas, but also in the ability to use the schemas efficiently to impose order on complex sound inputs.

Evidence suggests that infants have access to many of the same segregation cues that adults use. For example, infants appear to be able to use frequency, timbre, and amplitude differences to segregate a complex sound into separate “streams” (e.g., Fassbender 1993; Smith and Trainor 2011). However, it also appears that infants and children need larger frequency cues to perform this task (e.g., Sussman et al. 2007). Recently, informational masking paradigms have been used as a model of auditory stream segregation (Durlach et al. 2003; Leibold, Chap. 5). Developmental studies of informational masking indicate that by school age, children are able to use some cues (e.g., temporal cues, Leibold and Neff 2007; Leibold and Bonino 2009) as well as adults do. Additional studies delineating the cues that infants and children use in stream segregation and establishing the limits of children’s ability to use these cues are needed.

The studies just described paint a rather simple picture: Infants are able to use some cues to segregate auditory streams, and their ability to do so is progressively refined as they grow older and more experienced. The results of other studies of the development of auditory scene analysis suggest that the situation is more complicated. For example, apparently conflicting results have been reported with respect to

children's ability to use concurrent visual information to segregate one auditory stream from others, as when visual information from a speaker's face improves the ability to follow speech in a noisy background. Wightman et al. (2006) found that 6- to 9-year-old children were unable to use visual information to help them segregate one voice from another. This is a curious finding, because it has been known for some time that even young infants are sensitive to the correspondence between visual and auditory speech, at least for simple speech stimuli (e.g., Kuhl and Meltzoff 1982). Moreover, Hollich et al. (2005) found that 7.5-month-olds were able to segregate a target word from running speech at lower signal-to-noise ratio when the target word was accompanied by synchronous visual information. Importantly, a video of the talker's face and a generic trace on the screen that moved up and down with the target word were equally effective in promoting infants' segregation of the target word. The stimuli used in the Wightman et al. and the Hollich et al. studies, however, were quite different. Wightman et al. used the Coordinate Response Measure (CRM) procedure (e.g., Brungart et al. 2001): The target sentence and the distracter sentence were always of the form "Ready <call sign> go to <color> <number> now." The listeners reported the color and number named by the talker who used a specified call sign. In this situation, the target and distracter sentences are temporally similar, making it difficult to segregate the sentence on that basis. The difference in talker fundamental frequency is probably the most salient segregation cue. In the Hollich et al. study, infants' recognition of a repeated word presented in a background of monotone speech was tested. In this case, multiple segregation cues, including temporal cues, are available.

A similar pattern of results is seen in studies of children's use of spatial cues to segregate auditory streams. Hall et al. (2005) and Wightman et al. (2003) found that, for adults and older children, a masker that produced significant informational masking when it was presented in the same ear as the target produced little or no masking when it was presented to the contralateral ear. However, young children still exhibited significant informational masking with a contralateral masker. Similarly, using the CRM procedure, Wightman and Kistler (2005) found that whereas a contralateral masker had little effect on adults' or older children's performance, it caused significant amounts of masking for young children. The results of these studies documenting children's difficulty in using spatial (or at least lateralized) information to segregate speech streams contrast with those of Litovsky and her colleagues (Litovsky 2005; Johnstone and Litovsky 2006; Garadat and Litovsky 2007), who report that children as young as 3 years of age demonstrate significantly less masking of a target word presented at 0° azimuth when maskers were presented at 90° rather than at 0° azimuth. Johnstone et al., moreover, found that children were better able to use spatial cues than spectrotemporal differences between the target word and masker (e.g., speech vs. noise or speech vs. reversed speech). These two groups of studies differ, again, in the availability of temporal segregation cues, but also in the operational definition of "spatial": In studies reporting that young children do not use spatial cues to segregation, sounds are presented under earphones and the role of spatial information is assessed by comparing performance when the masker and target are presented to the same ear to that when the masker and target are

presented to opposite ears. In studies reporting that young children can use spatial cues to segregation, sounds are presented in free field and the role of spatial information is assessed by comparing performance when the target and masker come from the same location in space to that when masker and target come from difference locations in space. It is possible that dichotic presentation is not functionally equivalent to a difference in spatial location to infants and children (e.g., Friedman and Pastore 1977).

Interestingly, the studies reporting that children have difficulty using spatial cues also limited the temporal segregation cues, while those reporting that children can use spatial cues provide substantial temporal segregation cues. One conclusion might be that adults and children approach the problem of auditory scene analysis quite differently. Adults seem to have access to many cues. When one cue is salient, the addition of other cues may not improve performance. However, when some cues are limited, adults use whatever cues are available to segregate the auditory streams. Children, on the other hand, may depend heavily on one or a few cues and experience difficulty when those cues are not available, even when they have access to other cues. However, when the favored cue is available, children may still be able to benefit from the availability of other cues. In other words, the additivity of cues is very different in children and adults.

Once an auditory scene has been reconstructed, a listener's task becomes selecting a relevant object for further processing. The latter process is frequently called selective attention. As discussed by Leibold (Chap. 5), selective attention is believed to become more selective between infancy and adulthood, but as Leibold points out, it is often difficult to distinguish between a failure of auditory scene analysis and a failure of selective attention. Moreover, when listeners are asked to report a target among similar distracters, energetic masking may also be at work. The additivity of different types of masking has not been addressed extensively in adult listeners (but see Neff et al. 1988) and has not been addressed at all in developing listeners.

When the listener's task is to identify a speech target in speech distracters, one approach to disentangling these effects has been to analyze error patterns: If the listener has either not heard or not segregated the target from the distracters, then errors should be random, unrelated to either target or distracters. However, if the failure is at the selective attention stage, then listeners are likely to report items from the distracters when they make errors. Wightman and colleagues have used this approach extensively in analyses of children's performance in the CRM procedure. Wightman and Kistler (2005) analyzed error patterns of 4- to 16-year-olds and of adults to identify conditions in which performance was clearly limited by selective attention. Although even the oldest children did not perform as well as adults in the CRM task, with ipsilateral, contralateral, or both types of distracters, the effect of target-to-distracter ratio was similar across age groups, and error patterns suggested that ignoring the ipsilateral distracter was the limiting factor for all ages. Younger children, however, appeared unable to use level difference cues to segregate target from distracter, and in some conditions, had difficulty ignoring a distracter in the contralateral ear. Thus, younger children may exhibit immature auditory scene analysis as well as immature selective attention.

Event-related potentials represent another potentially fruitful approach to understanding how and why selective attention changes with age. A study by Bartgis et al. (2003) illustrates this approach: Children at 5, 7, and 9 years were tested in a dichotic listening task, and the P3 response was recorded to the target sound when it was in the attended ear and when it was in the unattended ear. Performance in the task improved with age, but more importantly, older children's P3 amplitude was greater for the attended target than for the unattended. In the 5-year-olds, there was no difference in response amplitude between the two conditions. Such a result makes a strong case that immature selective attention contributed to the 5-year-olds' poor performance. Whether age-related changes in event-related potentials (Eggermont and Moore, Chap. 3) will limit the extension of such an approach into infancy is not clear.

2.3 *Effects of Experience*

The role of experience with sound is an elemental concept in the field of development, and it is treated in several of the chapters in this volume (Eggermont and Moore, Chap. 3; Litovsky, Chap. 6; Panneton and Newman, Chap. 7; Trainor and Unrau, Chap. 8; Eisenberg et al., Chap. 9). The goal here is to identify some important observations in these chapters that have broader implications for the study of auditory development.

Eggermont and Moore (Chap. 3) present a developmental framework, in which the primary developmental event is maturation of brain stem circuitry during infancy, maturation of thalamocortical connections to the deep layers of auditory cortex between 6 months and 5 years of age, and the gradual maturation of intracortical and other connections in the superficial layers of auditory cortex between 5 and 12 years. Eggermont and Moore also discuss the effects of auditory deprivation on each of these developmental periods, based on observations of children who become deaf at different ages and who receive cochlear implants at different ages. The results of these studies are interesting in that they indicate that it is not always possible to predict when experience with sound is critical, relative to the period during which a particular system typically undergoes the most pronounced age-related change.

One of the more interesting findings reported in this literature is that although the development of brain stem circuitry reflected in the electric auditory brain stem response (EABR) depends on experience with sound, the time window during which such experience is effective is quite long. For example, Thai-Van et al. (2007) reported that in children with early-onset hearing loss who were implanted between 1 year and 12 years of age, EABR latencies developed along the normal trajectory, independent of age at implantation. This suggests that some auditory pathways retain plasticity in the absence of stimulation for quite a long time beyond the age at which they would normally mature.

In the case of the development of the thalamocortical pathway and the deep layers of primary auditory cortex, experience with sound is also necessary for development, but within a more restricted time window. For example, Sharma et al. (2002) found

that in deaf children who were implanted before 3.5 years of age, cortical P1 latencies developed along a normal trajectory, but that in children implanted after 7 years of age, asymptotic P1 latencies were longer than normal. In this case, experience throughout the period during which the pathway typically develops appears to be critical to the outcome.

Finally, Eggermont and Ponton (2003) reported that children who experienced a period of deafness of at least 3 years within the first 6–8 years of life never developed an N1 response, which is believed to be generated in the later developing superficial layers of auditory cortex. In the normal system, N1 emerges only toward the end of this period. Thus, experience with sound appears to be necessary in this case before the period during which the pathway normally undergoes the greatest developmental change.

There are, however, many limitations of this framework. For example, in Chap. 6, Litovsky discusses her work on spatial hearing in individuals who use cochlear implants. Among her findings are that prelingually deaf adults who use cochlear implants are sensitive to interaural level differences, but not to interaural time differences. Psychophysical data such as these do not tell us which parts of the brain are affected by auditory deprivation, but because interaural differences are calculated in brain stem nuclei, these results suggest that it may be a mistake to think of the brain stem as a unit. The pathways that generate the ABR do not encompass all of the brain stem pathways, and thus a normal ABR does not necessarily mean that all pathways are functioning normally.

Another limitation is related to the level of detail in our understanding of the effects of deprivation on auditory perception, reflected in both the broad strokes version of neural processing provided by evoked potentials and the relatively gross measures of auditory capacities typically used to assess the effects of hearing loss on auditory development. Consider the studies discussed by Eisenberg et al. (Chap. 9) that show that hearing loss early in life can be associated with deficits in speech understanding. Whereas some of the studies described by Eisenberg et al. tested phonetic discrimination, others tested word or sentence recognition, in quiet or in noise. However, as Panneton and Newman describe in Chap. 7, infants and children develop the ability not only to discriminate phonetic contrasts or to recognize words, but also to segment words from running speech and a variety of other abilities. It is important to consider auditory development in hearing-impaired children in a more molecular way to understand all the repercussions and complications that stem from degraded auditory input. Although the lack of audibility is clearly a major problem faced by hearing-impaired children, the broader consequences of limited acoustic information on mature speech perception must also be considered.

A final issue is that the current framework, based as it is on studies of individuals who have had little or no auditory input before activation of a cochlear implant, has limited applicability to children with mild-to-moderate hearing loss. Although it is clear from studies of plasticity in adult animals that such hearing loss can lead to reorganization of the auditory pathway (e.g., Seki and Eggermont 2002), it is less clear how such reorganization will limit perceptual development in prelingually hearing-impaired children. A lack of input may prevent the development of integrative cortical processes, but will partial or reduced input have similar effects?

3 Other Issues in the Study of Auditory Development

3.1 *Psychoacoustics, Speech, and Music*

A question generally not addressed is how development of psychoacoustic abilities is related to the development of speech and music perception. It is reasonable to ask, for example, how it is that infants who have relatively high pure-tone thresholds and poor high-frequency resolution are able to discriminate phonetic contrasts and melodies. One answer is that both speech and music perception can be fairly accurate, even when the auditory input is somewhat degraded, particularly when the experimenter or the parent takes pains to make sure that the sound is audible to an infant. This view is consistent with the observations of Nozza (1987) that infants' speech discrimination is more sensitive to reductions in sound level than is adults'. Another less investigated possibility is that because natural acoustic signals are redundant, containing many distinctive cues, infants may be discriminating the sounds like adults are, but on the basis of different information. Nittrouer and her colleagues (Nittrouer and Boothroyd 1990; Nittrouer 2004, 2005) have shown that this is the case in preschool children. However, the cues that infants use to distinguish naturally occurring sounds may be constrained by their auditory sensitivity. Such interactions between sensory coding and complex perception in infant listeners may be an interesting topic of future studies.

Another topic that has generated interest recently in developmental psychoacoustics is listening under difficult, or complex, conditions. Buss et al. (Chap. 4) discuss several cases in which children seem adult-like in their ability to process some acoustic cues, but then have difficulty processing the same cue if the situation is more complex. Litovsky (Chap. 6) notes that children are able to discriminate sound locations accurately in some conditions but have difficulty when "reverberation" is present. Leibold (Chap. 5) describes many situations in which young children have proportionately greater difficulty than adults in processing sounds, such as in randomly varying backgrounds or with limited segregation cues. An interest in the same general question has arisen in parallel in developmental speech perception. Panneton and Newman devote a section of Chap. 7 to the development of speech perception in nonoptimal listening conditions. There have been few points of intersection between developmental psychoacoustics and speech perception on this topic, although it is likely that the mechanisms involved in the development of listening in difficult conditions are the same whether the signal is a tone or speech. It may behoove those interested in speech perception to consider informational masking and those interested in psychoacoustics to consider the role of social factors in auditory development generally.

3.2 *Auditory Development Is a Prolonged Process*

For many years, the study of the development of speech perception, music perception, and psychoacoustic abilities has been limited to children younger than about 7 or 8 years of age. Of course, 7- or 8-year-olds' audition is adult-like in many

respects: Their absolute and masked sensitivity is mature; they have no difficulty processing the phonetic contrasts of their native language or segmenting words from running speech or learning new words; and they are sensitive to the properties of the music they hear, including ambient musical scales and key membership.

At the same time, adults' ability to process speech and music rapidly, efficiently, and accurately is extremely resistant to various sorts of degradation. The fact that speech, in particular, is an "overlearned" signal is thought to contribute to this resistance. A question for developmentalists is how long it takes for speech to become overlearned. A few recent studies suggest that as late as adolescence, speech processing is not as resilient as it is in adulthood. Teenagers appear to be less able to cope, for example, with the elimination of some contrastive cues in a speech identification task (Hazan and Barrett 2000). As Eggermont and Moore (Chap. 3) point out, however, integrative auditory cortical structures and responses continue to develop into adolescence. It is indeed the case that auditory development is a prolonged process. The perceptual consequences of later development have yet to be fully explored.

4 Summary

In the last 40 years, tremendous gains have been made in our understanding of human auditory development. A consistent story of auditory development is emerging from the study of psychoacoustic, speech, and music processing. This story can now be viewed from the framework of a developing auditory nervous system, and studies of atypical development have contributed greatly to an appreciation of the role of experience with sound in auditory development. Future work in this area will provide a deeper understanding of the precise mechanisms involved in development, and will benefit from greater interaction across subfields and a longer perspective on the time course of development.

Acknowledgments Preparation of this chapter was supported by funding from NIDCD, R01 DC00396.

References

- Abdala, C., & Slinger, Y. S. (1996). The development of cochlear frequency resolution in the human auditory system. *Ear and Hearing*, 17(5), 374–385.
- Adamson, C. L., Reid, M. A., Mo, Z. L., Bowne-English, J., & Davis, R. L. (2002). Firing features and potassium channel content of murine spiral ganglion neurons vary with cochlear location. *Journal of Comparative Neurology*, 447(4), 331–350.
- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *Journal of the American Academy of Audiology*, 21(9), 575–585.

- Ashmead, D. H., Clifton, R. K., & Perris, E. E. (1987). Precision of auditory localization in human infants. *Developmental Psychology*, 23(5), 641–647.
- Ashmead, D. H., Davis, D., Whalen, T., & Odom, R. (1991). Sound localization and sensitivity to interaural time differences in human infants. *Child Development*, 62(6), 1211–1226.
- Bartgis, J., Lilly, A. R., & Thomas, D. G. (2003). Event-related potential and behavioral measures of attention in 5-, 7-, and 9-year-olds. *Journal of Genetic Psychology*, 130(3), 311–335.
- Bertoncini, J. (1993). Infants' perception of speech units: Primary representation capacities. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 249–257). Dordrecht: Kluwer.
- Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology [General]*, 117(1), 21–33.
- Bertoncini, J., Serniclaes, W., & Lorenzi, C. (2009). Discrimination of speech sounds based upon temporal envelope versus fine structure cues in 5- to 7-year-old children. *Journal of Speech Language and Hearing Research*, 52, 682–695.
- Bibas, A., Liang, J., Michaels, L., & Wright, A. (2000). The development of the stria vascularis in the human foetus. *Clinical Otolaryngology*, 25(2), 126–129.
- Bibas, A. G., Xenellis, J., Michaels, L., Anagnostopoulou, S., Ferekidis, E., & Wright, A. (2008). Temporal bone study of development of the organ of Corti: Correlation between auditory function and anatomical structure. *Journal of Laryngology and Otology*, 122(4), 336–342.
- Bredberg, G. (1968). Cellular pattern and nerve supply of the human organ of Corti. *Acta Oto-Laryngologica Supplementum*, 236.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, 110(5), 2527–2538.
- Buss, E., Hall, J. W., Iii, & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *Journal of the Acoustical Society of America*, 120(5), 2777–2788.
- Clarkson, M. G., & Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of the Acoustical Society of America*, 77, 1521–1528.
- Clifton, R. K. (1992). The development of spatial hearing in human infants. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 135–157). Washington, DC: American Psychological Association.
- Clifton, R. K., Morrongiello, B., Kulig, J., & Dowd, J. (1981). Auditory localization of the newborn infant: Its relevance for cortical development. *Child Development*, 52, 833–838.
- Clifton, R. K., Gwiazda, J., Bauer, J., Clarkson, M., & Held, R. (1988). Growth in head size during infancy: Implications for sound localization. *Developmental Psychology*, 24, 477–483.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*, 114(1), 368–379.
- Eggermont, J. J. (1991). Frequency dependent maturation of the cochlea and brainstem evoked potentials. *Acta Oto-Laryngologica (Stockholm)*, 111, 220–224.
- Eggermont, J. J., Brown, D. K., Ponton, C. W., & Kimberley, B. P. (1996). Comparison of distortion product otoacoustic emission (DPOAE) and auditory brainstem response (ABR) traveling wave delay measurements suggests frequency-specific synapse maturation. *Ear and Hearing*, 17, 386–394.
- Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: Correlations with changes in structure and speech perception. *Acta Oto-Laryngologica*, 123(2), 249–252.
- Fassbender, C. (1993). *Auditory grouping and segregation processes in infancy*. Norderstedt, Germany: Kaste Verlag.

- Folsom, R. C., & Wynne, M. K. (1987). Auditory brain stem responses from human adults and infants: Wave V tuning curves. *Journal of the Acoustical Society of America*, 81, 412–417.
- Friedman, C., & Pastore, R. E. (1977). Effects of lateralization on selective and divided attention. *Journal of the Acoustical Society of America*, 62, S1–S2.
- Garadat, S. N., & Litovsky, R. Y. (2007). Speech intelligibility in free field: Spatial unmasking in preschool children. *Journal of the Acoustical Society of America*, 121(2), 1047–1055.
- Gilley, P. M., Sharma, A., Dorman, M., & Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clinical Neurophysiology*, 116(3), 648–657.
- Gomes, H., Dunn, M., Ritter, W., Kurtzberg, D., Brattson, A., Kreuzer, J. A., & Vaughan, H. G. (2001). Spatiotemporal maturation of the central and lateral N1 components to tones. *Developmental Brain Research*, 129(2), 147–155.
- Hall, J. W., & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal-modulation transfer-function. *Journal of the Acoustical Society of America*, 96(1), 150–154.
- Hall, J. W., Buss, E., & Grose, J. H. (2005). Informational masking release in children and adults. *Journal of the Acoustical Society of America*, 118(3), 1605–1613.
- Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, 28(4), 377–396.
- Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, 76(3), 598–613.
- Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *Journal of the Acoustical Society of America*, 120(4), 2177–2189.
- Jusczyk, P. W., Pisoni, D. B., Walley, A., & Murray, J. (1980). Discrimination of relative time of two-component tones by infants. *Journal of the Acoustical Society of America*, 67, 262–270.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *Journal of the Acoustical Society of America*, 94, 2617–2638.
- Keefe, D. H., Burns, E. M., Bulen, J. C., & Campbell, S. L. (1994). Pressure transfer function from the diffuse field to the human infant ear canal. *Journal of the Acoustical Society of America*, 95, 355–371.
- Kinney, H. C., Brody, B. A., Kloman, A. S., & Gilles, F. H. (1988). Sequence of central nervous system myelination in human infancy 2. Patterns of myelination in autopsied infants. *Journal of Neuropathology and Experimental Neurology*, 47(3), 217–234.
- Kuhl, P. K. (1991). Human adults and human infants show a “perceptual effect” for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93–107.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138–1140.
- Leibold, L. J., & Bonino, A. Y. (2009). Release from informational masking in children: Effect of multiple signal bursts. *Journal of the Acoustical Society of America*, 125(4), 2200–2208.
- Leibold, L. J., & Neff, D. L. (2007). Effects of masker-spectral variability and masker fringes in children and adults. *Journal of the Acoustical Society of America*, 121(6), 3666–3676.
- Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*, 117(5), 3091–3099.
- Marin-Padilla, M., & Marin-Padilla, T. M. (1982). Origin, prenatal development and structural organization of layer I of the human cerebral (motor) cortex—a Golgi study. *Anatomy and Embryology*, 164(2), 161–206.
- Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing*, 3(6), 301–308.
- Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America*, 54, 610–619.
- Moore, J. K., & Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2(4), 297–311.

- Morrongiello, B. A. (1988). Infants' localization of sounds along the horizontal axis: Estimates of minimum audible angle. *Developmental Psychology*, 24, 8–13.
- Neff, D. L., Jesteadt, W., & Callaghan, B. P. (1988). Combined masking under conditions of high uncertainty. *Journal of the Acoustical Society of America*, 83, S33.
- Newman, R. S., & Evers, S. (2007). The effect of talker familiarity on stream segregation. *Journal of Phonetics*, 35(1), 85–103.
- Ng, M. (2000). Postnatal maturation of the human endolymphatic sac. *Laryngoscope*, 110(9), 1452–1456.
- Nittrouer, S. (2004). The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults. *Journal of the Acoustical Society of America*, 115(4), 1777–1790.
- Nittrouer, S. (2005). Age-related differences in weighting and masking of two cues to word-final stop voicing in noise. *Journal of the Acoustical Society of America*, 118(2), 1072–1088.
- Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*, 87, 2705–2715.
- Nozza, R. J. (1987). Infant speech-sound discrimination testing: Effects of stimulus intensity and procedural model on measures of performance. *Journal of the Acoustical Society of America*, 81(6), 1928–1939.
- Nozza, R. J., & Wilson, W. R. (1984). Masked and unmasked pure-tone thresholds of infants and adults: Development of auditory frequency selectivity and sensitivity. *Journal of Speech and Hearing Research*, 27, 613–622.
- Okabe, K. S., Tanaka, S., Hamada, H., Miura, T., & Funai, H. (1988). Acoustic impedance measured on normal ears of children. *Journal of the Acoustical Society of Japan*, 9, 287–294.
- Olsho, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior & Development*, 8, 371–384.
- Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *Journal of the Acoustical Society of America*, 82, 454–464.
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America*, 84(4), 1316–1324.
- Paetau, R., Ahonen, A., Salonen, O., & Sams, M. (1995). Auditory-evoked magnetic fields to tones and pseudowords in healthy children and adults. *Journal of Clinical Neurophysiology*, 12(2), 177–185.
- Pasman, J. W., Rotteveel, J. J., Degraaf, R., Maassen, B., & Notermans, S. L. H. (1991). Detectability of auditory evoked response components in preterm infants. *Early Human Development*, 26(2), 129–141.
- Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology*, 18(15), 1124–1128.
- Pujol, R., & Lavigne-Rebillard, M. (1995). Sensory and neural structures in the developing human cochlea. *International Journal of Pediatric Otorhinolaryngology*, 32(Supplement), S177–182.
- Pujol, J., Soriano-Mas, C., Ortiz, H., Sebastian-Galles, N., Losilla, J. M., & Deus, J. (2006). Myelination of language-related areas in the developing brain. *Neurology*, 66(3), 339–343.
- Rotteveel, J. J., de Graaf, R., Colon, E. J., Stegeman, D. F., & Visco, Y. M. (1987). The maturation of the central auditory conduction in preterm infants until three months post term. II. The auditory brainstem responses (ABRs). *Hearing Research*, 26, 21–35.
- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America*, 86, 1733–1742.
- Schneider, B. A., Morrongiello, B. A., & Trehub, S. E. (1990). The size of the critical band in infants, children, and adults. *Journal of Experimental Psychology [Human Perception and Performance]*, 16, 642–652.
- Seki, S., & Eggermont, J. J. (2002). Changes in cat primary auditory cortex after minor-to-moderate pure-tone induced hearing loss. *Hearing Research*, 173(1–2), 172–186.

- Sininger, Y. S., Abdala, C., & Cone-Wesson, B. (1997). Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response. *Hearing Research*, 104(1–2), 1–22.
- Sinnott, J. M., Pisoni, D. B., & Aslin, R. M. (1983). A comparison of pure tone auditory thresholds in human infants and adults. *Infant Behavior & Development*, 6, 3–17.
- Smith, N. A., & Trainor, L. J. (2011). Auditory stream segregation improves infants' selective attention to target tones amid distracters. *Infancy*, 16, doi: 10.1111/j.1532-7078.2011.00067.x.
- Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, 61, 632–652.
- Sussman, E., Wong, R., Horvath, J., Winkler, I., & Wang, W. (2007). The development of the perceptual organization of sound by frequency separation in 5–11-year-old children. *Hearing Research*, 225, 117–127.
- Thai-Van, H., Coma, S., Boutitie, F., Disant, F., Truy, E., & Collet, L. (2007). The pattern of auditory brainstem response wave V maturation in cochlear-implanted children. *Clinical Neurophysiology*, 118(3), 676–689.
- Tharpe, A. M., & Ashmead, D. H. (2001). A longitudinal investigation of infant auditory sensitivity. *American Journal of Audiology*, 10(2), 104–112.
- Trehub, S. E., Schneider, B. A., Morrengiello, B. A., & Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *Journal of Experimental Child Psychology*, 46, 273–285.
- Trehub, S. E., Schneider, B. A., & Henderson, J. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustical Society of America*, 98, 2532–2541.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.
- Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development*, 63, 260–272.
- Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*, 118(5), 3164–3176.
- Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development*, 60, 611–624.
- Wightman, F., Callahan, M. R., Lutfi, R. A., Kistler, D. J., & Oh, E. (2003). Children's detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*, 113(6), 3297–3305.
- Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *Journal of the Acoustical Society of America*, 119(6), 3940–3949.
- Yakovlev, P. I., & Lecours, A.-R. (1967). The myelogenetic cycles of regional maturation of the brain. In A. Minkowski (Ed.), *Regional development of the brain in early life* (pp. 3–70). Oxford: Blackwell.

Chapter 2

Morphological and Functional Ear Development

Carolina Abdala and Douglas H. Keefe

*The aim of argument, or of discussion, should not be victory,
but progress.*

—Joseph Joubert

1 Introduction

The development of peripheral auditory function in humans has been observed and documented using a variety of investigative tools. Because these tools must all be noninvasive in nature, they are indirect and, therefore, somewhat imprecise probes of function. Measured function at one level of the peripheral auditory system is undoubtedly influenced by the functional status of other parts of the system. Thus, the most effective way to define and document the physiology and developmental course of the human auditory system is to consider and integrate findings, with an acute awareness of the limitations of each assay and the relationship among results. To make this treatise a reasonable endeavor, only *human* auditory developmental data are presented and discussed, although when it elucidates a pattern common to humans, mammalian development in general may be considered.

Section 2 of the chapter deals with the most peripheral segments of the auditory system, the outer and middle ear, and with acoustical measurements of the functioning

C. Abdala (✉)
Division of Communication & Auditory Neuroscience, House Research Institute,
Los Angeles, CA, USA
e-mail: cabdala@hei.org

D.H. Keefe
Boys Town National Research Hospital, Omaha, NE, USA
e-mail: Douglas.Keefe@boystown.org

of the ear canal and middle ear in adults and infants. Section 3 examines what is known about the development of human cochlear function focusing on otoacoustic emission (OAE) measurements in human adults and infants. Following these descriptions of functional development, Sect. 4 attempts to interpret the findings and consider their possible *sources*, speculating, when possible, about the implications for development of hearing. The authors of this chapter sometimes have very different views on the sources of maturational differences in the described data. To the extent that readers become informed on diverse viewpoints concerning current issues in the field, this may be of some benefit. The diverse viewpoints presented here should encourage readers to access additional literature en route to formulating their own well-reasoned opinions.

2 Development of External- and Middle-Ear Responses

2.1 *Maturational Changes in Anatomy*

The anatomy of the human ear canal continues to mature after full-term birth. The canal is straighter and shorter in infants than in adults (Northern and Downs 1984), for example, a length of 1.68 cm in newborns, which is roughly two thirds of that in adults (Crelin 1973). The canal wall of the newborn has no bony portion and consists of a thin, compliant layer of cartilage (Anson and Donaldson 1981), and the tympanic ring does not completely develop until age 2 years (Saunders et al. 1983). The adult canal has a bony wall in its inner two thirds and soft-tissue wall in the outer one third, whereas the canal in newborns is almost completely surrounded by soft tissue (McLellan and Webb 1957). The ear-canal diameter and length each increases from birth through the oldest test age of 24 months (Keefe et al. 1993). Human temporal bone data show maturational changes in the middle ear during childhood. The tympanic-membrane plane relative to the central axis of the ear canal is more horizontal in the newborn, with a more adult-like orientation by age 3 years (Ikui et al. 1997).

The tympanic membrane develops embryologically as a structure composed of ectodermal, mesodermal, and endodermal layers. The outer layer is similar to the epidermis of skin; the lamina propria is composed of a matrix and two layers of type II collagen fibers, and the thin inner lamina mucosal layer with columnar cells forms a boundary of the tympanic cavity (Lim 1970). The collagen fibers of the mucosal layer provide mechanical stiffness to the tympanic membrane (Qi et al. 2008). The thickness of the adult tympanic membrane has broad variability across its surface as well as across subjects; it ranges from 0.04 to 0.12 mm in a central region of the pars tensa (Kuypers et al. 2006). The tympanic membrane in newborns is thicker than in adults, with a thickness ranging from 0.4 to 0.7 mm in the posterior–superior region, 0.7–1.5 mm in the umbo region, and 0.1–0.25 mm in the posterior–inferior, anterior–superior, and anterior–inferior regions (Ruah et al. 1991). Qi et al. (2008) confirmed that the measured thickness of the tympanic membrane in the temporal bone of a 22-day-old infant was in the range of the Ruah et al. data.

The volume of the middle-ear cavity increases postnatally until the late teenage years (Eby and Nadol 1986); this growth may change the ossicular orientation and influence the mechanical functioning of the middle ear. A smaller middle-ear cavity volume also increases middle-ear stiffness at an ear-canal location just in front of the tympanic membrane, because the membrane motion drives against the volume compliance. The volume of the middle-ear cavity includes contributions from the tympanic cavity, the aditus ad antrum, the mastoid antrum, and mastoid air cells. The tympanic-cavity volume is approximately 640 mm^3 in adults and 452 mm^3 in 3-month-old infants (Ikui et al. 2000), and approximately 330 mm^3 in a 22-day-old. The mastoid process begins to develop approximately 1 year after birth. The ossicles of the middle ear are completely formed around the sixth month of fetal life (Crelin 1973) and the middle-ear muscles are then fully developed (Saunders et al. 1983). The distance between the stapes footplate and the tympanic membrane in adults is larger than that of infants through 6 postnatal months (Eby and Nadol 1986). Dimensions of ear-canal and tympanic-membrane components in the adult and newborn ear (with the latter based on temporal-bone data from a 22-day-old) are tabulated in Qi et al. (2006).

2.2 *Maturational Changes in Function*

2.2.1 **Diffuse-Field Effects**

Under free-field listening conditions in the adult human ear, the transfer-function level between sound pressure at the ear drum relative to a location near the opening of the ear canal is close to 0 dB at low frequencies, and has a resonant boost in pressure near 2.7 kHz. This boost is accompanied by a more broadly tuned concha resonance near 4.5–5 kHz that leads to an overall gain between 2 and 7 kHz in the adult ear (Shaw and Teranishi 1968). The ear-canal resonance frequency near 3.0 kHz in children ages 2–12 years remains slightly higher than adult values (Dempster and Mackenzie 1990). Thus, maturation of ear-canal function extends at least to the onset of puberty.

The diffuse-field absorption cross-section A_D quantifies the ability of the external ear to collect sound power from free-field sources and of the middle ear to absorb this power (Rosowski et al. 1988; Shaw 1988). For a sound source sufficiently far from a listener, the sound field in a room approximates a diffuse field averaging out the directional properties of sound. Neglecting power loss at the canal walls, A_D is the ratio of the sound power absorbed by the middle ear relative to the diffuse-field power density in the room. It has units of area; its level is defined as $10 \log_{10} A_D$ with 0 dB at 1 cm^2 .

In measurements on adults and in infants of age 1–24 months, this transfer function increased with increasing frequency at approximately 6 dB per octave at low frequencies up to a maximum level at a resonance frequency that varied with age (see Fig. 2.1). The frequency of maximum response increased with decreasing age from 2.5 kHz in adults to 5 kHz in 1-month-olds. A_D was larger in adults than in

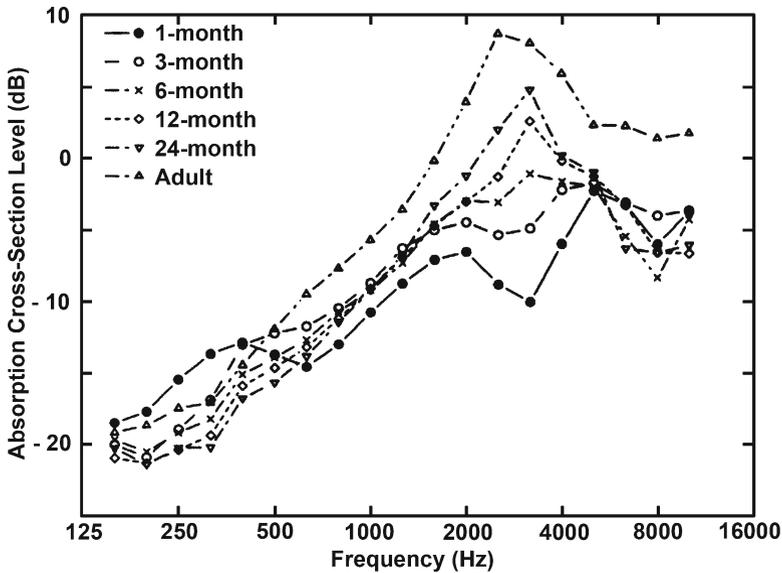


Fig. 2.1 The mean diffuse-field absorption cross-section A_D is plotted as a function of frequency for adults and infants in the listed age groups [This figure was originally published in Keefe et al. (1994)]

infants, and decreased by approximately 10 dB with increasing age (Keefe et al. 1994). Thus, external and middle ears are more efficient at collecting sound energy in adults compared to infants. In relative terms across frequency, the efficiency of infant ears is maximal at a higher frequency than in adults. This is partially explained by maturational growth in the cross-sectional area of the ear canal area. A larger area leads to more efficiency in collecting and absorbing sound power, and thus a larger A_D . Because ear-canal area increases with age, the ability of the ear to absorb sound power improves with increasing age, whether sound is presented in the free field or via some type of coupler measurement.

2.2.2 Maturation Factors of Ear-Canal Acoustics

Many hearing experiments and clinical measurements are performed using a probe microphone and sound source in a probe inserted in a leak-free manner into the ear canal. Such a procedure removes acoustic influences of the outer ear and torso, and thus all directional hearing cues. The only remaining outer-ear effect is ear-canal acoustics, which vary with age.

Using median measurements of the ear canal area for infants (A) and for adults ($A_0 = 58 \text{ mm}^2$) (Keefe and Abdala 2007), the distributions of the area measured in groups of infants were normalized to the median adult area in terms of the relative area level (in dB) by $\Delta L_a = 20 \log_{10}(A_0/A)$.

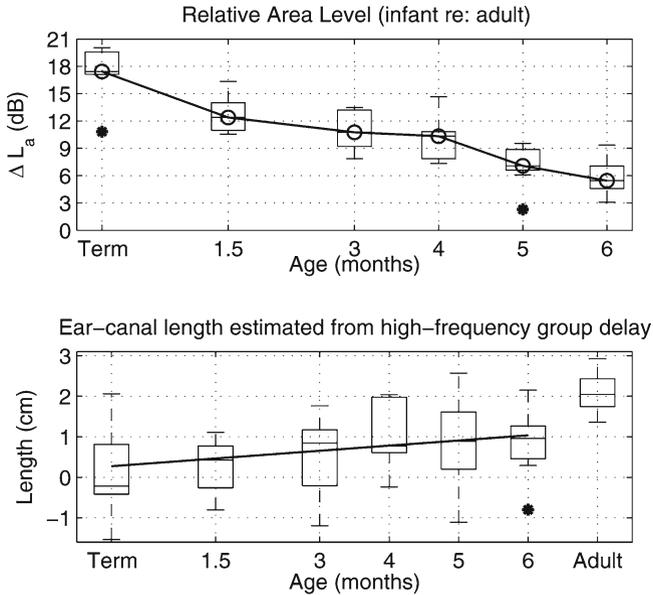


Fig. 2.2 The group norms for relative area level ΔL_a (*top*) and ear-canal length (*bottom*) are plotted for each age group using box-and-whisker plots, i.e., the top and bottom lines of each box show the interquartile range; the *horizontal line* within the box shows the median; and any outlier data that lie outside the range of near outliers, which is defined by the breadth of the “whiskers” above and below each box, are plotted as asterisks. In addition, the medians of ΔL_a are also plotted by open circles connected by a *solid line* (*top*). A linear regression line for infant ear-canal length is plotted in the *bottom panel*

These are plotted in the top panel of Fig. 2.2, with outlier responses shown by asterisks. The median of the relative area level was 17 dB larger for term infants than for adults, and decreased to 5 dB at 6 months.

The median ear-canal length based on Keefe and Abdala (2007) increased with increasing age from newborns to adults (bottom panel of Fig. 2.2). Negative length values in the tails of the infant distributions result from small errors in the estimates. A regression line for ear-canal length across the infant age groups shows that length increased from 0.3 cm to 1.1 cm with increasing age to 6 months. This contrasts with the 2-cm length in adult canals between the probe and the tympanic membrane. These length estimates correspond to the geometric length minus the insertion depth of the probe assembly.

The relatively large interquartile ranges of ear-canal length and area in Fig. 2.2 show the critical importance of individual variability at each age. Measurements at ages up to 24 months show that neither the ear-canal length nor the area is mature at 24 months (Keefe et al. 1993), and it is likely ear-canal growth continues at a slower rate until the onset of puberty. From the standpoint of air-conducted sound, the ear canal acts as a filter whose magnitude and phase characteristics change ever more

slowly from infancy through adulthood. The frequency response of the filter influences the representation of sound encoded neurally, which would affect subjective properties such as loudness and timbre.

2.2.3 Absorbance in the Ear Canal

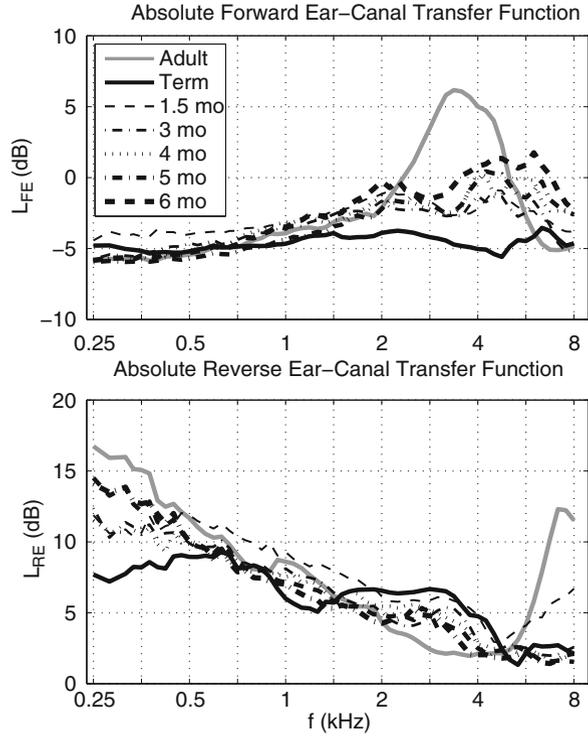
Aural absorbance is a dimensionless acoustic transfer function that assesses how efficiently the middle ear, and the ear canal to the extent that any wall loss is present, absorbs power (Feeney and Keefe 2010). It has particular relevance for hearing experiments and clinical measurements performed using a sound source within the ear canal. Absorbance varies between 0 and 1, with a value of 1 when the ear absorbs all the power from the forward-traveling sound in the ear canal, and a value of 0 when the ear absorbs no power (so that all the power is reflected at the tympanic membrane). The mean absorbance has a maximum in the range from 2 to 4 kHz for adults and infants of age 1.5 months and older. In term newborns, the peak occurs slightly below 2 kHz. This contrasts with the peak frequency in the diffuse-field absorption cross section (Fig. 2.1), which varied with frequency and occurred at higher frequencies for younger infants. Thus, maturational differences in ear-canal and middle-ear function vary depending on whether the sound source is in a free field or delivered into the ear-canal using a small insert probe.

Because the wave nature of sound produces significant effects above 1 kHz in the adult ear canal, it is not simple to decompose ear-canal and middle-ear function into independent “black boxes” as the location of the sound source is varied. Models incorporating wave effects are constructed to perform such a decomposition using measured data. This is why absorbance is so useful, because its value does not vary with probe location within the canal. In contrast, acoustic admittance, which is another aural acoustic transfer function that is widely used in clinical tympanometry, varies with probe location within the ear canal. The ability to decompose an admittance tympanogram into an ear-canal volumetric component and a middle-ear component is only possible at frequencies near 226 Hz (in the absence of a model that takes into account the wave nature of sound and the motion of the ear-canal walls). Switching from admittance to absorbance allows a study of middle-ear function over a much broader frequency range (i.e., up to 8 kHz). This has clinical translational significance as further described in Feeney and Keefe.

2.2.4 Forward Ear-Canal Transfer Function

A forward ear-canal transfer function level L_{FE} describes the forward transmission of sound from the ear-canal probe to the tympanic membrane. L_{FE} is defined as the magnitude (level in dB) of the ratio of the forward ear-canal pressure to the total pressure at the probe (Keefe and Abdala 2007). L_{FE} is relevant to the interpretation of any hearing experiment in which sound is presented within the ear canal. It describes how differences in ear-canal length and cross-sectional area influence the

Fig. 2.3 The mean L_{FE} (top) and L_{RE} (bottom) are plotted versus frequency for adult and infant age groups



forward-traveling pressure wave that is directed toward the eardrum in terms of the total pressure measured by the probe microphone, and includes standing-wave effects in the ear canal. It also depends on the processes of sound absorption and reflection at the tympanic membrane. L_{FE} varies systematically with age as a result of the maturational changes in the acoustical functioning of the ear canal and middle ear.

Wideband measurements of L_{FE} in infants and adults show a prominent peak in the adult ear close to 3.4 kHz, but no major peak in infants (see top panel of Fig. 2.3). This emphasizes the importance of standing waves in the adult ear canal between 3 and 4 kHz, and confirms the relatively flat response in the shorter ear canals of infants. The fact that the ear-canal length is about twice as long in the adult ear as in the 6-month-old helps explain differences in L_{FE} between adults and infants in Fig. 2.3. The pronounced maximum L_{FE} in adults is a standing-wave effect in the ear canal that is much reduced in infants.

The general pattern suggested by Fig. 2.3 is that wideband acoustic transfer functions of the ear canal and middle ear are immature in infants. These functions remain non-adult-like until at least 11 years of age (Okabe et al. 1988). This long period of maturation throughout childhood is consistent with the slow maturation of some components of ear-canal and middle-ear anatomy described in Sect. 2.1. Maturation differences in these transfer functions affect stimulus levels used in hearing development experiments.

2.2.5 Reverse Ear-Canal Transfer Function

An evoked OAE measurement presents forward-directed sound stimuli, which elicit a cochlear-generated signal that is reverse-transmitted back to the ear canal. Unlike other acoustic variables used in hearing experiments, OAEs depend not only on L_{FE} but also on a reverse ear-canal transfer function level (L_{RE}). L_{RE} is defined in terms of the magnitude of the ratio of the total pressure measured at the ear-canal probe microphone to the reverse-transmitted pressure in the canal adjacent to the tympanic membrane. This transfer-function level is plotted for adult and infant age groups in the bottom panel of Fig. 2.3 based on results in Keefe and Abdala (2007). Aside from excess variability across age below approximately 0.5 kHz, where OAE responses are rarely measured, L_{RE} in infants varies within a few decibels relative to adult levels between 0.5 and 6 kHz. This contrasts with the large difference in L_{FE} between 2.8 and 4 kHz that was observed between infants and adults. Above 6 kHz, L_{RE} is much larger in adult than infant ears, which acts to attenuate the level of high-frequency OAEs in infants relative to adults.

2.2.6 Forward Middle-Ear Transfer Function

At any particular frequency, a forward middle-ear transfer function level L_{FM} describes the forward transmission of sound through the middle ear from the tympanic membrane to the cochlear vestibule. It is defined in terms of the ratio of the forward-directed pressure difference at the basal end of the cochlea to the ear-canal forward pressure just in front of the tympanic membrane. The total forward transfer function from the ear canal to the cochlear base is $L_F = L_{FE} + L_{FM}$. For any hearing experiment, knowledge of L_F would describe the transformation of stimulus level associated with ear-canal and middle-ear function. Although L_{FE} was straightforward to evaluate based on measurements of aural acoustic transfer functions (see top panel of Fig. 2.3), a measurement of L_{FM} , and thus L_F , would require highly invasive procedures that are impossible to perform in human subjects for ethical reasons.

From the standpoint of understanding hearing development, it is important to understand maturational differences in how sound is conducted through the ear canal and middle ear, or, in the general case of free-field listening, through the outer ear and middle ear. Although this conductive pathway functions as a linear system (in the absence of middle-ear muscle reflex effects that occur at relatively high sound levels), it, nonetheless, may have a profound effect on cochlear nonlinearity, and thus on the neural encoding of sound. As considered in detail later in this chapter, cochlear function has a compressive nonlinearity. For example, this means that a 1-dB increase in the sound pressure level in the ear canal produces an increase of less than 1 dB in the amplitude of vibration of the basilar membrane. If the forward conduction of sound through the outer and middle ear would differ between infant and adult ears, then the effective input to the cochlea would also differ. Any such change in input level would modify the “set point” on the compressive nonlinearity of cochlear mechanics. Thus, any maturational effects of linear middle-ear functioning become intermixed with effects of cochlear nonlinearity.

An OAE experiment differs from other hearing experiments in that the OAE is measured in the ear canal. As further described in Sect. 4.2, a combination of non-invasive measurements can reveal how L_{FM} differs between infant and adult ears.

2.2.7 Newborn Ear-Canal Wall Motion Effects

The increased compliance of the ear-canal wall in infants compared to older children and adults has functional consequences. Video-otoscopy results show that air pressures close to 300 daPa produced changes (with significant intersubject variability) of up to 70% in ear-canal diameter in some 1- to 5-day-olds, 10% changes or less in infants between 31 and 56 days, and no detectable change beyond 56 days (Holte et al. 1991). It is reasonable that variations in acoustic pressure within the canal might also produce variations in wall displacement. A model of ear-canal wall motion in infants was developed using a viscoelastic model of the mechanics of how the wall moves in response to pressure changes (Keefe et al. 1993). Properties of aural acoustic transfer functions below 1.2 kHz, which were observed in 1- to 3-month-old infant responses but not in responses from older children and adults, were predicted by this model. Whereas sound introduced into the ear canal of an older child or adult acts on the air enclosed within the ear canal and the tympanic membrane, sound in the neonatal ear canal also acts “in parallel” to drive ear-canal wall motion. Finite-element modeling of the newborn ear canal at frequencies between 226 Hz and 1 kHz confirms that sound presented in the ear canal produces a sinusoidal volume change of the ear canal (Qi et al. 2006). A finite-element model of a 22-day-old newborn middle ear (Qi et al. 2006) predicts that the ear-canal wall volume displacement may be larger than the volume displacement of the tympanic membrane. This is a confounding factor in interpreting standard 226-Hz probe tone tympanograms in full-term infant ears in the first 2 or 3 months of development.

2.2.8 Acoustic Reflex Effects

High-level sound presented in the ear canal can elicit a contraction of the stapedius muscle of the middle ear to produce a middle-ear muscle reflex (MEMR). This change can be detected in terms of a measured change in level or phase of a probe sound in the ear canal. An acoustic reflex test is influenced by ear-canal, middle-ear, cochlear, and afferent and efferent neural function. A wideband test of the acoustic reflex was used to measure thresholds in 80 ears of normal-hearing adult and 375 ears of newborn infants with normal hearing based on having distortion product OAE (DPOAE) responses within a normal range (Keefe et al. 2010). Although the MEMR is typically assumed to extend to frequencies below 2.8 kHz, reflex shifts were observed at frequencies up to 8 kHz. These observed shifts may also include contributions from the medial olivocochlear efferent system in addition to the MEMR.

MEMR thresholds were measured using a broadband noise activator to elicit the reflex. Thresholds were quantified in terms of the minimum activator SPL at which a reflex shift was observed. Infant thresholds based on in-the-ear SPL were only

2 dB higher than adult thresholds. Inasmuch as this 2-dB difference was small compared to variability in thresholds within each age group, the MEMR thresholds using broadband noise did not appear to differ between newborn infants and adults. Interpreting this outcome is complicated by procedural differences between infant and adult ears in how a reflex shift was classified as present or absent, and by the fact that the forward transmission of sound in the broadband noise activator and probe signal varied across frequency as well as across age.

3 Development of Human Cochlear Function

3.1 *Developmental Changes in Anatomy*

General mammalian cochlear development proceeds from base to apex (Bredberg 1968; Pujol and Hilding 1973). This gradient applies both to general architecture as well as cellular and subcellular mechanisms reviewed below (Pujol et al. 1991). Around the fifth embryological week in humans, the otocyst is formed by a cartilaginous matrix (Sanchez-Fernandez et al. 1983). At 9 weeks, three cochlear coils are apparent with a well-differentiated otic capsule and a septum separating the coils. Around this time, the process of hair cell differentiation begins. Afferent fibers, and possibly efferent, are entering the otocyst and forming what appears to be targeted, specific synapses well before hair cells are fully differentiated. At this point in development, transmission electron microscopy shows presumptive hair cells in the epithelium that can be distinguished from supporting cells by cytoplasmic content and afferent fibers at their base (Pujol and Lavigne-Rebillard 1985).

By weeks 10 and 11, a more definitive differentiation of inner (IHCs) and outer hair cells (OHCs) begins with IHC preceding OHC development. Stereocilia starts to form on IHCs and about a week later, on OHCs. This sequence exemplifies the second developmental gradient noted in the cochlea: IHC differentiation, morphological development and innervation all precede comparable processes in the OHC. Once stereocilia begin to form, microvilli disappear quickly from the cuticular plate and the remaining cilia are organized in three or four parallel rows, graded in length. Lateral and tip links are observed as soon as stereociliary organization is established (Lavigne-Rebillard and Pujol 1986; Lim and Rudea 1992). Hair cell length is changing at this time, along with stereocilia length. IHC length is similar throughout cochlea; however, OHC length depends on the cell's position along the basilar membrane and increases from base to apex (Pujol et al. 1992). By weeks 20 and 21, there are adult-like stereocilia bundles on both types of hair cells (Lavigne-Rebillard and Pujol 1986; 1987). The general shape of the OHC remains immature much longer than IHC shape (Pujol and Hilding 1973). Both types of cells are organized in rows but OHCs are less regimented and more chaotic in their organization compared to highly organized IHCs. At 20 fetal weeks, OHCs in the apical cochlea are sparse and their organization is particularly chaotic, with irregular spacing and

uneven development of ciliary bundles (Tanaka et al. 1979). At some point in development, supernumerary sensory cells that are generally not seen in the mature system are present in the developing cochlea (Bredberg 1968; Igarishi 1980). The functional significance of these extra sensory cells is not well understood.

The protracted course for OHC development in mammals is loosely correlated with changes in functional properties of the cochlea such as increasing sensitivity, frequency tuning, and shifts in the place code (Pujol et al. 1980; Romand 1987). For example, there is an accumulation of actin in the cuticular plate (Raphael et al. 1987), redistribution of mitochondrial and endoplasmic reticulum, and the formation of intricate subsurface cisternae preceding OHC motility observed in vitro. OHC somatic motility is thought to power the cochlear amplifier, which augments tuning and auditory sensitivity. The human OHC appears morphologically mature, including postsynaptic specialization, some time in the third trimester, though this final stage has not been well delineated. It is not clear whether subtle morphological immaturities remain after term birth (Pujol et al. 1998).

The physical features of the basilar membrane have not been characterized during fetal and early postnatal life. Imaging of delicate membranes in postmortem temporal bones may not provide an effective means of characterizing features such as basilar membrane mass and/or its stiffness gradient. Supernumerary hair cells have been identified in fetal and neonatal tissue and could affect mass of the membrane, as could differences in its dimensions. At present, the maturational time course for the dimensions and material properties of the basilar membrane in the human auditory system is not known.

Afferent VIIIth nerve fibers stop spreading and focus on the base of the existing sensory cells once the hair cells are fully differentiated. They innervate IHCs around the 11th or 12th fetal week (Lavigne-Rebillard and Pujol 1986). Vesiculated *efferent* endings are observed in the inner spiral sulcus below IHCs around the 14th week (Pujol and Lavigne-Rebillard 1985). The subsequent axosomatic synapses that form between large medial efferent fibers and IHCs are transient and may represent a transitional period (Simmons 2002). The formation of more permanent, classic efferent axosomatic synapses with OHCs occurs sometime between 20 and 30 weeks (Lavigne-Rebillard and Pujol 1990). These synapses are most scarce at the apex of the cochlea and continue to show immaturities, such as incompletely developed postsynaptic features and small presynaptic varicosities into the third trimester and possibly beyond.

The morphological and functional development of the cochlea does not determine the onset of hearing. Auditory nerve fibers must be able to conduct sound-evoked impulses to produce hearing. This milestone occurs sometime around 25–27 weeks as defined by in utero studies of fetal movement or heart rate (Birnholtz and Benacerraf 1983) and 27–28 weeks as defined by onset of auditory brain stem responses (ABRs) that reflect averaged, synchronous nerve and brain stem potentials (Galambos and Hecox 1978). Although present by 27 fetal weeks, the ABR is not mature when initially observed and requires approximately the first year of postnatal life to become adult-like in morphology and speed of transmission (i.e., latency) as detailed in Eggermont and Moore, Chap. 3.

3.2 *Maturational Changes in Function*

3.2.1 **Basic OAE Characteristics**

Otoacoustic emissions provide a unique, noninvasive window into human cochlear function through which auditory peripheral maturation can be studied. OAEs are preneural cochlear responses that can be easily and noninvasively recorded with a sensitive microphone and earphone assembly placed at the entrance of the ear canal, making them ideal measures to study human cochlear function.

Early studies reported spontaneous OAEs (SOAEs) to be present in infants and adults with equal prevalence, around 65% (Strickland et al. 1985; Burns et al. 1992). However, later reports found higher prevalence in premature and term newborns, ranging from 82% to 90% (Morlet et al. 1995; Abdala 1996). By all accounts, newborns have higher level SOAEs than adults (Burns et al. 1992). It is not unusual to observe a newborn SOAE in the range of 20–25 dB SPL compared to a more modest 5–10 dB SPL signal in a young adult. In addition, newborns also have more numerous SOAEs per ear than adults, although they show adult-like sex and ear trends: females and right ears have more SOAEs.

Approximately 90–100% of normal-hearing adults and newborns exhibit click-evoked OAEs (CEOAEs) (Bonfils et al. 1989). Prematurely born neonates in the NICU are reported to have slightly lower prevalence rates of approximately 80–90%, especially if they are tested within the first 3 days of birth, most likely due to debris and fluid that drains from the external auditory canal and middle-ear space in the first 72 h (Stevens et al. 1990). Newborns have higher level CEOAEs in general than adults and adolescents; response levels decrease with increasing age (Norton and Widen 1990; Prieve 1992).

Distortion product OAEs (DPOAEs) are also present in 90–100% of normal-hearing adults and newborns (Lonsbury-Martin et al. 1990; Bonfils et al. 1992). Neonates tend to have slightly higher DPOAE levels (~2–5 dB) than adults, generally in the low- and mid-frequency range (Lasky et al. 1992; Smurzynski et al. 1993). Prematurely born infants tested at 33–36 weeks postconceptional age have mildly reduced DPOAE levels compared to their term-born counterparts; however, levels increase with age through 40 weeks, at which time they are comparable to those observed in term neonates (Smurzynski 1994; Abdala et al. 2008). Between birth and 6 months of age, there is little change in OAE level, though OAEs continue to show greater magnitude than adult emissions well into childhood (Prieve et al. 1997a, b).

Spontaneous OAEs are centered near 2.3 kHz for adults and 3.5 kHz for newborns (Strickland et al. 1985; Burns et al. 1992; Abdala 1996). Some studies have indicated a consistent upward shift of SOAE frequency (in absence of amplitude shift) as a function of postconceptional age in prematurely born infants (Briennesse et al. 1997). Other studies have reported stable SOAE frequency through 24 months of age in human infants (Burns et al. 1994). It appears that the bias toward high frequencies disappears sometime between birth and preadolescence (Strickland et al. 1985). Evoked OAEs from neonates generally show an extended high-frequency

spectral content compared to those from adults. Both CEOAE and DPOAE spectra are flatter and exhibit a wider bandwidth than the adult response (Lasky et al. 1992; Abdala et al. 2008).

3.2.2 OAE Generation Mechanisms

Background

Otoacoustic emissions are generally present in the normal mammalian cochlea if the cochlear amplifier, driven by OHC somatic motility, is intact and functional. Thus, OAEs reflect normal OHC physiology and cochlear amplifier function. With little exception, when OAEs are absent (as long as the conductive system is normal), cochlear function can be judged to be abnormal. In the last 10 years, an updated taxonomy has been developed that differentiates among OAE types by their source region on the basilar membrane (Knight and Kemp 2001) and by their distinct generation mechanisms (Shera and Guinan 1999). Investigating these generation mechanisms in the developing human auditory system may provide a more targeted study of how and when distinct aspects of human cochlear function mature.

It is hypothesized that different OAE types reflect at least two cochlear processes (Shera and Guinan 1999). DPOAEs are a “mixed” type emission, encompassing both cochlear processes, and thus best elucidating this framework of OAE generation. In addition, DPOAEs are typically recorded with moderate level stimulus tones in newborns, making it possible to achieve adequate signal-to-noise ratio (SNR). Other OAEs, such as stimulus frequency OAEs (SFOAEs), may be less confounded by mixed sources but require a low-level stimulus. SFOAE measurements have only recently been reported in human newborns (Kalluri et al. 2011). Recent theory, supported by strong experimental evidence, contends that DPOAEs comprise two fundamentally different mechanisms: (1) OHC-based intermodulation distortion, which is an intrinsically nonlinear process (generated at the overlap region between primary tones) and (2) linear coherent reflection initiated near the DPOAE frequency by place-fixed irregularities or “roughness” along the length of the basilar membrane (Zweig and Shera 1995; Shera and Guinan 1999).

The phase behavior of each of the two DPOAE components differs markedly (Talmadge et al. 1998; Shera and Guinan 1999). For distortion or “wave-fixed” components generated at the overlap of the f_1, f_2 primary tones (and recorded with a fixed f_2/f_1 ratio), phase is relatively invariant over most of the frequency range. The distortion source is induced by the wave (in this case, the interaction of the two primary tones) thus, the source moves with the wave as stimuli are swept in frequency. Since DPOAE phase is referenced to the phases of the primary tones, as long as the f_2/f_1 ratio is kept constant, DPOAE phase should be constant. In contrast, the reflection or “place-fixed” component generated at $2f_1 - f_2$ is linked to irregularities along the membrane; hence, the corresponding reflection phase shifts with these irregularities because they are fixed in position. This produces rapidly rotating phase, and consequently a steep OAE phase slope (Shera and Guinan 1999).

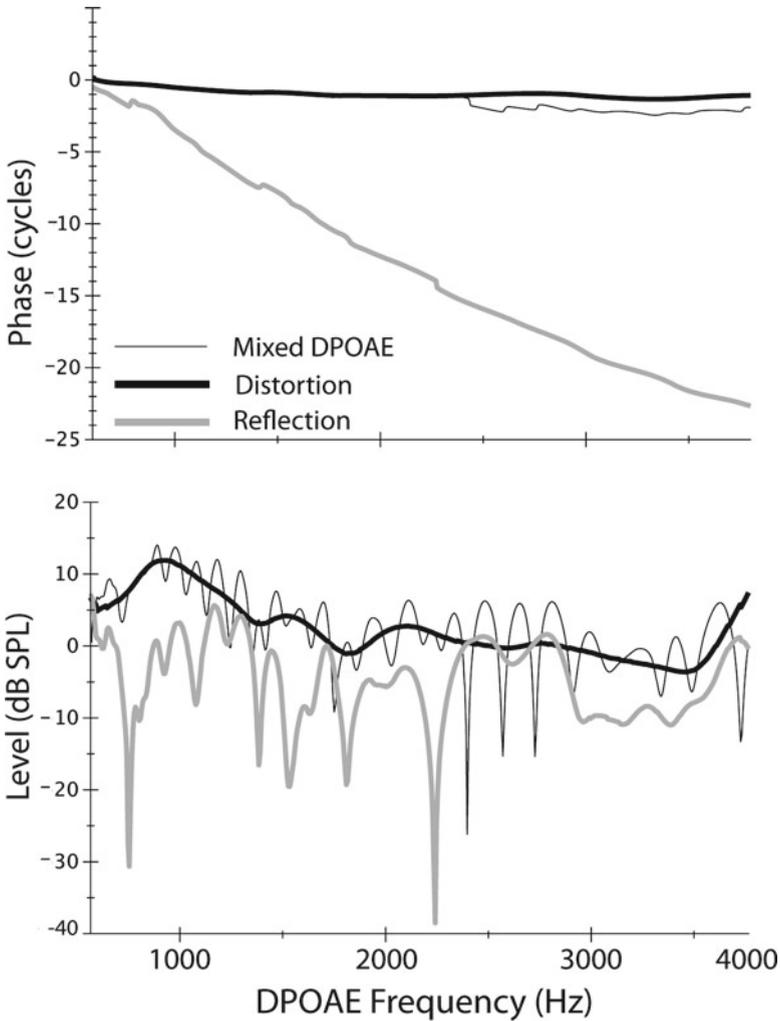


Fig. 2.4 Magnitude and phase output from an inverse fast Fourier transform (IFFT) conducted on DPOAE data from one adult subject. Phase of the ear canal or “mixed” DPOAE (*thin black line*), the distortion component (*thick black*) and the reflection component (*gray line*) are shown in the *upper panel*. Corresponding mixed DPOAE and component magnitude are shown in the *lower panel*

Output from an inverse fast Fourier transform (IFFT) conducted to separate the two components is shown in Fig. 2.4. The upper panel of Fig. 2.4 shows the phase behavior from the “mixed” ear canal DPOAE as well as each of the two DPOAE components after separation by a time windowing technique. The DPOAE measured at the microphone in the ear canal represents the vector sum of these components, which combine constructively and destructively. It is the difference in phase rotation of the components that produces DPOAE fine structure shown in Fig. 2.5.

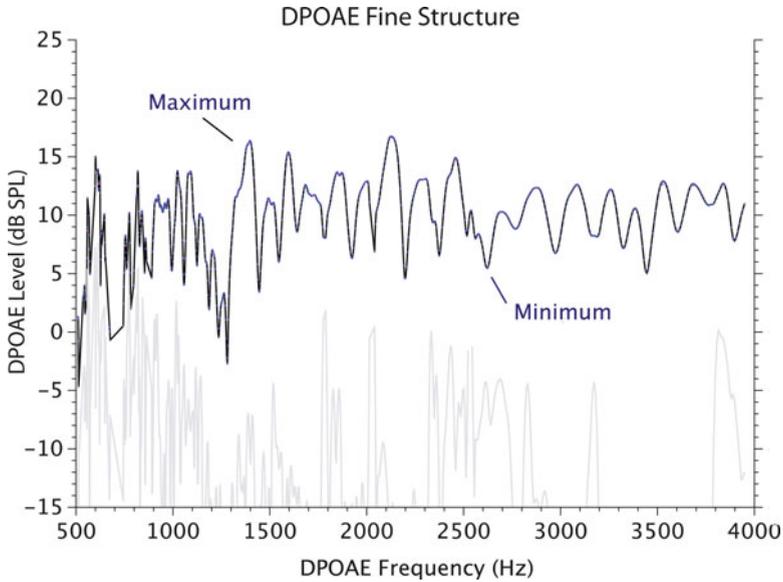


Fig. 2.5 When recorded with high resolution, DPOAE level exhibits alternating peaks (maxima) and valleys (minima) or *fine structure*, reflecting interference between two DPOAE sources

When they are similar in magnitude and 180° out-of-phase, components cancel, producing dips or minima in fine structure; when they add while in-phase, DPOAE level is augmented as noted by peaks or maxima in fine structure.

Overall, at moderate and high stimulus levels, the nonlinear distortion component of the DPOAE is dominant in normal ears across most of the frequency range, as noted in the lower panel of Fig. 2.4. CEOAEs, SFOAEs, and SOAEs exhibit phase features consistent with linear reflection linked to fixed-position perturbations along the basilar membrane. There are several reviews written to describe this theoretical framework of OAE generation in detail and the reader is referred to these for additional information (Shera and Guinan 1999; Shera 2004; Shera and Abdala 2010).

Development

How can this emerging theoretical framework of OAE generation contribute to the study of peripheral auditory system development? If OAE components reflect distinct cochlear properties, do these distinct properties mature at different rates and does the relative contribution of each source-type to the OAE measured at the microphone change during infancy or childhood? Does OAE phase and its rotation as a function of frequency offer a clue about basilar membrane motion during development? To what extent do immaturities in outer- and middle-ear function contribute to the observed immaturities in OAE responses? These are intriguing questions that

are now being explored to better understand the maturation of cochlear function. As of yet, few studies have separated and individually analyzed DPOAE components in human infants to scrutinize targeted cochlear properties.

A recent series of experiments (Abdala and Dhar 2010; Abdala et al. 2011b) has begun to describe DPOAE fine structure, phase, and individual DPOAE components in human newborns. These investigations found more prevalent DPOAE fine structure and narrower spacing between oscillations in newborns compared to adults. An IFFT and time-windowing technique was applied to separate DPOAE components (see Fig. 2.4) and found that, although the distortion-source component of the DPOAE was similar in magnitude between adults and newborns, the reflection component was significantly larger in the infants. Enhanced newborn reflection, coexisting with adult-like distortion levels, prompts speculation that the peripheral properties underlying each component may have distinct maturational time courses. The implications of this finding are considered in Sect. 4. The ability to study two unique OAE components, each purported to represent distinct physiological processes within the cochlea, may allow for more targeted study of peripheral maturation. The need for further study notwithstanding, it may be possible to move beyond simple categorical distinctions of mature versus immature to more precisely characterize age-related (or with a bit of optimism, pathology-related) changes in targeted cochlear properties such as OHC nonlinearity, cochlear tuning and cochlear amplification.

3.2.3 DPOAE Phase and Basilar Membrane Motion

OAE phase may offer a unique glimpse into aspects of cochlear function. DPOAE phase is influenced by the material properties of the basilar membrane such as the spatially distributed stiffness, mass and damping. In gerbil, active processes (cochlear amplifier gain and sharp tuning) in the cochlear base are eliminated after death, whereas basilar membrane phase is minimally altered (Ren and Nuttall 2001). These findings suggest that active processes can be impacted independently from the more gross aspects of passive basilar membrane motion. Likewise, it may be possible to examine their maturational time courses somewhat independently. OAE phase may provide a tool for this purpose.

Some relatively early studies measured CEOAE and/or DPOAE onset latency and phase in an attempt to estimate cochlear travel time in human infants. However, these produced widely disparate results, which are now understood to be primarily due to methodological differences (Brown et al. 1994, 1995; Eggermont et al. 1996). As new theories of OAE generation have emerged, DPOAE phase has taken on renewed importance and more attention has been paid to its precise quantification and interpretation (Shera et al. 2000; Tubis et al. 2000). Two recent observations have reported DPOAE phase from human infants (Abdala and Dhar 2010; Abdala et al. 2011b). In the first of these studies, DPOAE phase was analyzed as group delay in adults and term newborns. Although both age groups showed a prolongation of group delay in the low frequencies (consistent with a break from cochlear

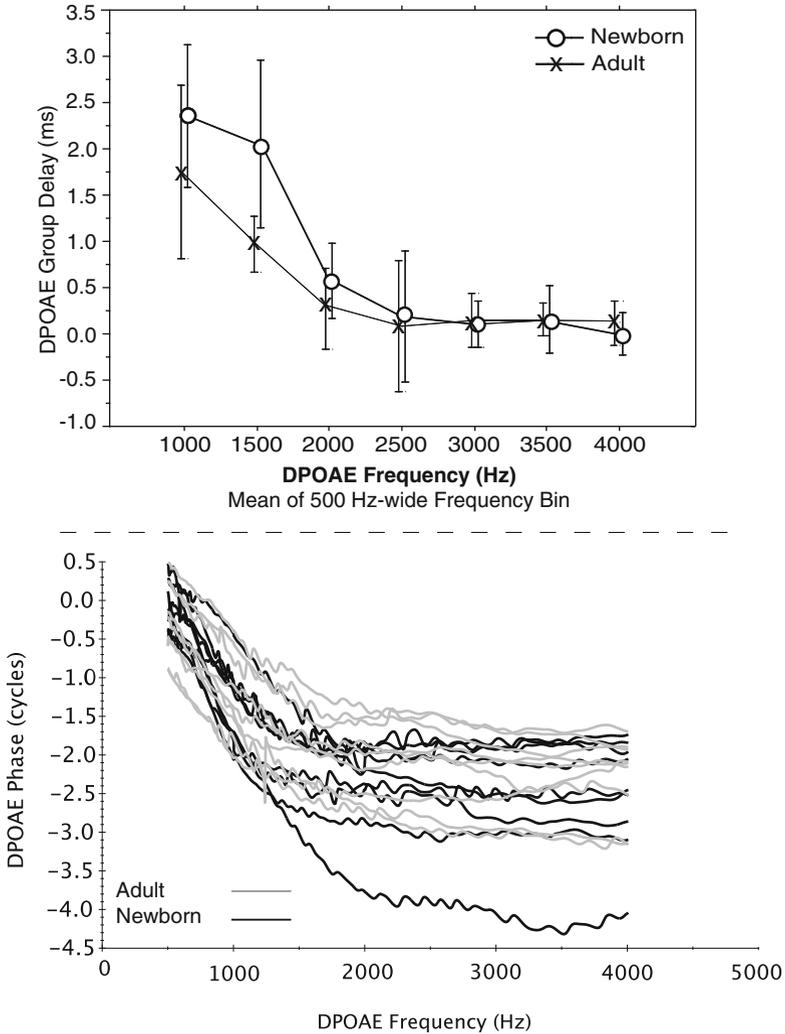


Fig. 2.6 The *upper panel* shows DPOAE mean group delay, ± 1 SD, for a group of newborns and young adults from Abdala and Dhar (2010); the *lower panel* displays individual DPOAE phase-frequency functions from a second, independent group of newborns and adults. Both figures show invariant phase (or group delay) through approximately 1,500 Hz and a relatively steep phase slope below this frequency

scaling in the apical half of the mammalian cochlea), the newborn delay was significantly more prolonged than the adult delay (upper panel of Fig. 2.6). A longer group delay represents a steeper DPOAE phase gradient.

A subsequent experiment extended these initial findings using a targeted, low-frequency protocol with extensive averaging to ensure adequate SNR and to further scrutinize apical cochlear function in newborns. Initially, a suppressor tone was

presented near the DPOAE frequency ($2f_1-f_2$) to ensure that the DPOAE was dominated by the distortion component (not the reflection source). The lower panel of Fig. 2.6 shows the resulting phase versus frequency functions for infant and adult subjects. The “break” frequency (elucidating the transition from invariant to steeply sloping phase) as well as the phase slope of the segment above and below this transition frequency were calculated. The break, thought to reflect the apical–basal demarcation, was centered near 1.4 kHz for both adults and newborns. However, consistent with the previous report, newborns exhibited significantly steeper slope of phase than adults in apical regions of the cochlea. The implications of this finding with respect to cochlear scaling are considered in Sect. 4.1.

3.2.4 Cochlear Response Growth

The relationship between a signal presented to the cochlea and the resulting cochlear output can be represented with an input/output (I/O) function. Although influenced by many complex factors, the OAE I/O function may provide insight into cochlear compressive nonlinearity. The typical adult OAE I/O function shows a relatively linear increase in amplitude with low stimulus levels and response saturation at mid- to high levels. In adults, click-evoked OAEs clearly show this pattern of nonlinear, compressive growth (mean slope=0.34) (Prieve 1992). There are no detailed data on CEOAE I/O functions in newborns and sparse data describing pediatric CEOAE I/O functions in general. The most comprehensive study included subjects aged 6 months through 17 years and reported that CEOAE I/O functions were roughly similar in shape and slope across all age groups (Prieve et al. 1997a).

DPOAE response growth has been fairly well characterized during development. Early reports described neonatal DPOAE I/O functions as nonmonotonic, with slope values ranging from 0.5 to 0.6 depending on frequency (Popelka et al. 1995). Others described I/O slope values that ranged from 0.25 at the lowest frequencies to 0.9 at the highest ($f_2=10$ kHz) (Lasky et al. 1992; Lasky 1998). More recent studies (Abdala and Keefe 2006; Abdala et al. 2007) reported that infant DPOAE I/O functions were displaced upward, consistent with higher DPOAE amplitude, and were more linear compared to adult functions. Up to 37% of I/O functions from premature newborns did not plateau, even at maximum stimulus levels of 85 dB SPL (Abdala 2000), whereas adults rarely showed functions without a saturating segment. Overall, adult I/O functions saturated around 65–70 dB SPL, whereas the infant functions that showed saturation plateaued at stimulus levels of 75–80 dB SPL. The DPOAE saturation threshold moved downward toward adult values throughout the premature period (33 weeks PCA) and during the first 6 postnatal months (Abdala 2003; Abdala et al. 2007). Figure 2.7 shows mean DPOAE I/O functions from infants tested at birth and again at 3, 4, 5, and 6 months of age. The source of age effects on the DPOAE I/O function appear to be related, not to cochlear immaturity as initially hypothesized, but rather to middle and ear canal immaturities (considered further in Sect. 4.1).

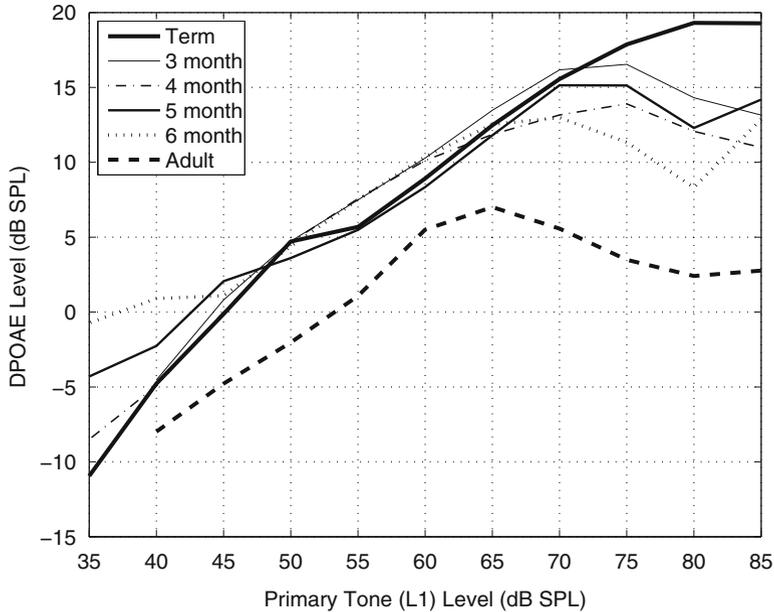


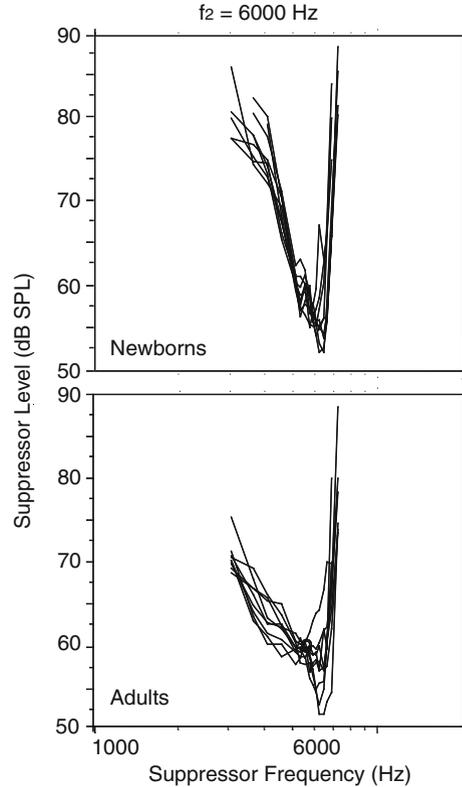
Fig. 2.7 Mean DPOAE input/output functions recorded in a group of adults and infants tested at birth and then at four additional postnatal ages as denoted in the key [This figure was originally published in Abdala and Keefe (2006)]

3.2.5 Development of Cochlear Frequency Selectivity

OAE ipsilateral suppression has been applied to the study of cochlear tuning in human infants. OAE amplitude is measured in the test ear at a fixed primary tone level and f_2/f_1 while ipsilateral tones are presented across a range of frequencies and levels to suppress the emission. Iso-response OAE suppression tuning curves (STCs) are generated by plotting the suppressor level required to produce a fixed-criterion amplitude reduction as a function of suppressor frequency. In the normal ear, STCs show morphology that is similar to auditory nerve frequency threshold curves and psychoacoustic tuning curves: a narrow tip centered slightly higher than the f_2 frequency, a steep high-frequency flank, and a shallower low-frequency flank with a tail-like segment for lower-frequency primary tones (Brown and Kemp 1984; Martin et al. 1987).

Early studies of DPOAE STCs generated with primary tones of $f_2=3$ kHz and 6 kHz (and a 6 dB suppression criteria) found roughly comparable tuning curve shape in term-born neonates and adults (Abdala et al. 1996). However, subsequent studies including larger groups of prematurely born as well as term-born infants, and more detailed quantification of STC features, found clear age differences in DPOAE suppression tuning between newborns and adults (Abdala 1998, 2001, 2003). Newborns showed *better*, narrower DPOAE suppression tuning with a steeper slope

Fig. 2.8 Individual DPOAE ipsilateral suppression tuning curves (6 dB suppression criterion) recorded at $f_2=6,000$ Hz from a group of 15 adults and 15 term-born neonates



on the low-frequency flank and a deeper, sharper tip than adults at $f_2=6$ kHz (see Fig. 2.8). At $f_2=1.5$ kHz, more subtle age differences were noted (e.g., the tuning curve width was narrower in newborns) but they were not as consistent, possibly because of elevated noise floors in the low-frequency range.

In lieu of generating iso-response STCs, plotting DPOAE level as a function of increasing suppressor level characterizes the growth of suppression across frequency. Tones on the low-frequency side of the probe ($<f_2$) produce linear or expansive suppression growth whereas high-frequency side suppressors ($>f_2$) produce suppression that is compressive and shallow (Abdala and Chatterjee 2003), as shown in Fig. 2.9. This pattern is consistent with both mechanical and neural responses observed in laboratory animals and closely mimics the two-tone suppression patterns recorded at various levels of the auditory system (Abbas and Sachs 1976; Delgutte 1990; Ruggero et al. 1992). It suggests that suppressors around or higher than f_2 are impacted by the expected compressive nonlinearity (just basal to the peak of the f_2 traveling wave), and thus are not effective suppressors as their level is increased. Conversely, suppressor tones lower in frequency than f_2 produce approximately linear reduction of the DPOAE with increases in level as the “tail” of the traveling wave effectively suppresses the more basal f_2 site. As noted in Fig. 2.9,

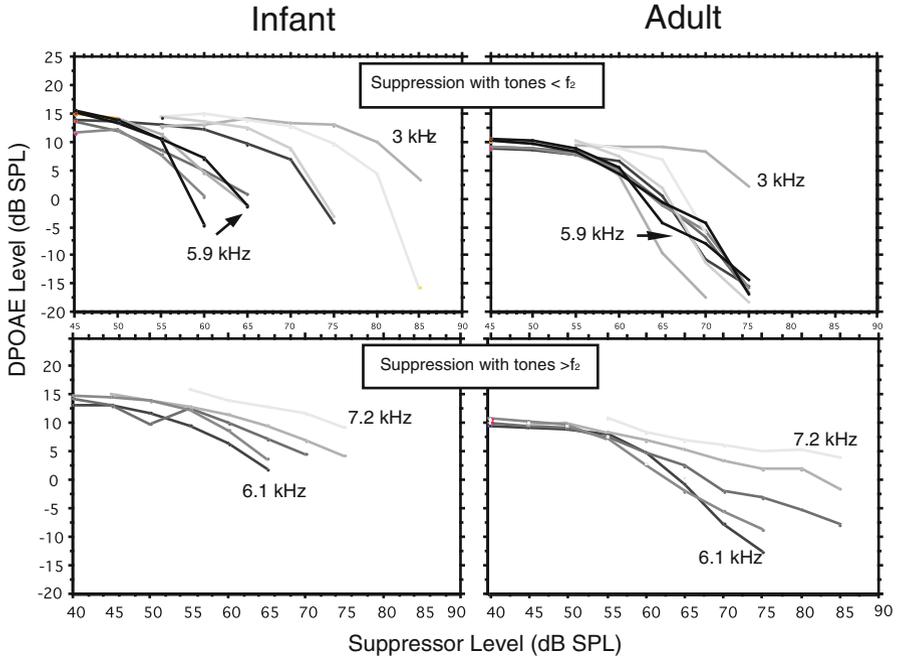


Fig. 2.9 DPOAE amplitude as a function of suppressor level ($f_2=6,000$ Hz; L1–L2=65–55 dB SPL) for one newborn on the *left* and one adult on the *right*. Suppressor tones with frequencies lower and higher than f_2 are presented in the *upper* and *lower panels*, respectively. Line shading roughly distinguishes suppressor frequency

newborns show elevated suppression thresholds and non-adultlike (more compressive) growth of suppression for suppressors $<f_2$. This pattern of suppression growth explicates the source of the steep low-frequency flank in newborn STCs.

A longitudinal study tracking DPOAE suppression tuning and suppression growth in a group of nine premature neonates over a 2-month period (33 weeks PCA through 38–40 weeks PCA) confirmed that the narrow tuning, steeper low-frequency flank and the sharper, deeper tip of the DPOAE STC at 6 kHz does not become adult-like by the equivalent of term birth (Abdala 2003). A second study (Abdala et al. 2007) extended this finding to the postnatal period and showed that STC width (Q_{10}), tip-to-tail measures, STC low-frequency slope and tip level remain immature, consistent with narrower and sharper tuning through at least 6 months of age, as noted in Fig. 2.10.

3.2.6 The Medial Olivocochlear (MOC) Reflex

No discussion of peripheral auditory development is complete without considering maturation of the descending, inhibitory efferent fibers that modulate cochlear function.

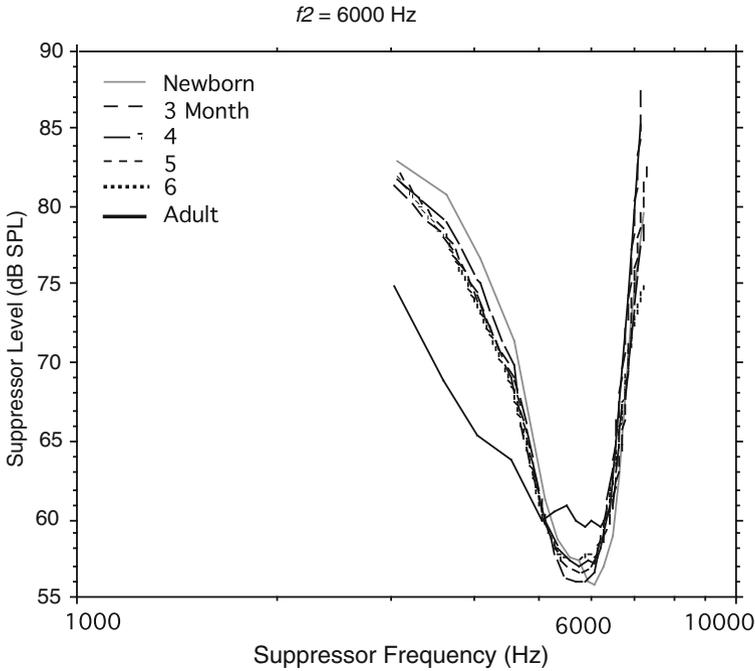


Fig. 2.10 Mean DPOAE suppression tuning curves ($f_2=6,000 \text{ Hz}$, $L1-L2=65-55 \text{ dB SPL}$) recorded from young adults and a group of infants tested initially at birth and then again at 3, 4, 5, and 6 months of age [This figure was originally published in Abdala et al. (2007)]

The MOC system is thought to enhance auditory perception in difficult listening conditions and provide a measure of protection from noise damage (Micheyl and Collet 1996; Maison and Liberman 2000). In laboratory animals, both visual and auditory inhibitory synapses have critical periods during development that influence later perception if disrupted (Walsh et al. 1998; Morales et al. 2002). In addition, as described in the previous section on maturation of cochlear anatomy, the medial efferent innervation of OHCs is one of the later developing processes in cochlear development, occurring sometime late in the third trimester and possibly into the perinatal or early postnatal period.

Early work with laboratory animals (Mountain 1980; Siegel and Kim 1982) confirmed that electrically stimulating the olivocochlear (OC) bundle in the brain stem alters cochlear output. Later work in humans showed that *acoustic* stimulation can evoke similar activation of OC fibers, producing reduction in OAE magnitude, presumably by hyperpolarizing OHCs (Puel and Rebillard 1990; Collet 1993). The *medial* portion of the OC tract (MOC) is strongly cholinergic and predominantly innervates OHCs through both crossed and uncrossed pathways (Fex 1962). A reduction in OAE amplitude when the MOC system is activated reflects MOC-induced inhibition (Guinan 2006). One typical paradigm used to probe MOC-mediated

regulation of cochlear function includes the presentation of broadband noise (BBN) in the contralateral ear while OAE level is monitored in the ipsilateral or test ear.

The contralateral MOC reflex has been studied in human newborns, though methodology varied significantly among studies. Given updated theories of OAE generation, it is clear that appropriate controls may not have been included and parameters were likely not optimized in these early studies (Morlet et al. 1993; Ryan and Piron 1994; Abdala et al. 1999). In general, CEOAE-based measures of the MOC reflex in newborns were reported to be immature in prematurely born infants, but adult-like by term birth (Morlet et al. 1993; Ryan and Piron 1994). DPOAE-based measures of the MOC reflex, likewise, found that adults and term-born infants exhibit comparable MOC effects at low and mid frequencies (Abdala et al. 1999). The magnitude of the reflex on average was 1.2 dB at $f_2 = 1.5$ kHz. Premature newborns, in contrast, did not show a significant MOC reflex; however, nearly half of their data included episodes of *increasing* DPOAE level (i.e., enhancement) produced by contralateral BBN. When enhancement values were eliminated, the magnitude of the MOC reflex overlapped for the three age groups suggesting that the *strength* of the reflex may be adult-like even during the premature period. Nevertheless, the increased prevalence of DPOAE level enhancement in prematurely born infants is noteworthy and may be related to interference between dual DPOAE sources.

The most current DPOAE-based protocols to measure the MOC reflex have taken the dual-mechanism model into account and calculate the reflex at fine structure peaks only or for the distortion- and reflection-source components separately (Abdala et al. 2009; Deeter et al. 2009). Consider the following sequence of events: (1) The DPOAE is recorded at a frequency where distortion and reflection components sum in the ear canal while 180° out-of-phase, producing cancellation (i.e., at a minimum frequency in fine structure); (2) the MOC reflex is activated by BBN and alters the reflection-source more than the distortion-source component, consequently shifting the phase relationship between them; (3) the shifted phase relationship releases phase cancellation, thus producing an abrupt and marked increase in DPOAE level. This DPOAE enhancement appears to be largely due to component interference. To address component interference (and its effects on measures of the MOC reflex), a recent investigation recorded the MOC reflex at DPOAE fine structure peak frequencies. This ensures that DP components were in-phase (Abdala et al. 2009). More than 90% of resulting MOC reflex values showed inhibition with this technique and overall mean MOC reflex values were in the range of 1.5–2.0 dB in young adult ears.

DPOAE-based MOC reflex protocols, with additional controls for component interference, phase- and magnitude-based reflex metrics and MEMR measures, are currently being applied to prematurely born and term newborns to assess the maturation of the MOC system (Abdala et al. 2011c). Preliminary data from 23 term-born neonates shows an overall mean MOC inhibitory reflex of 1.3 dB in newborns compared to 1.1 dB in adults. When measuring the MOC reflex as a vector difference (i.e. considering BBN-induced phase changes as well as magnitude effects) term newborns and adults showed mean MOC reflexes (expressed as a fraction of baseline DPOAE amplitude) of 0.21 and 0.16, respectively. These

preliminary results suggest robust, adult-like MOC reflexes at term birth; however, reflex *strength* may not be the only metric of maturity. It will be informative to assess the *efficiency* of the MOC reflex in providing enhanced signal in noise detection. These efforts are currently ongoing. The development of careful OAE-based measurement and analysis protocols for the MOC reflex may help disentangle the hypothesized sources of auditory peripheral immaturity in human newborns.

4 Interpretation of Findings

In the preceding section, non-adultlike OAE results from human infants were reported. The interpretation of these findings is not straightforward because the OAE is generated in the cochlea (and thus, sensitive to cochlear function) but is influenced by ear-canal and middle-ear function, as well as by descending medial efferent system function. The complexity of the task notwithstanding, the overarching objective is to explore what an infant hears. Thus, the following section attempts to place the reported findings into a cohesive, or at least interesting, explanatory framework for readers. Section 4.1 explains the OAE immaturities based on the possibility of cochlear immaturity (or non-adultlike function) during the perinatal and/or early postnatal period. Section 4.2 explains these immaturities in OAE responses based on the theory that cochlear function is mature at birth.

4.1 Cochlear Explanations

4.1.1 Active Processes

Adult–newborn differences in the prevalence, level, and number of SOAEs per ear have been observed. The global standing wave theory (Kemp 1979; Zweig 1991; Zweig and Shera 1995) predicts that SOAEs are standing wave resonances produced by multiple internal *reflections*, initiated by external sound or intrinsic physiological noise. Reflections occur off of irregularities on the basilar membrane and are “...self-sustaining when the total round trip power gain matches the energy loss (i.e., viscous damping and acoustic radiation in the ear canal) experienced en route” (Shera 2003, p. 245). Consistent with numerous and robust SOAEs in newborns, when the DPOAE is separated into its two components (distortion and reflection), the relative contribution of the reflection source is enhanced in newborns compared to adults. Thus, both SOAE and component-specific DPOAE findings are indicative of stronger cochlear reflection at birth.

The enhanced newborn OAE reflection component could be interpreted in several ways. It might indicate that the infant basilar membrane has more irregular or “rough” architecture and less regimented cellular organization, producing more robust back-scattering of energy; however, more roughness alone would not ensure

coherence of reflected wavelets. More likely, it indicates stronger cochlear amplification in newborns relative to adults. Reflection emissions are a low-level source, thought to mirror efficiency of the cochlear amplifier as they arise from the peak of the traveling wave. The coherent reflection model explains this type of OAE as reflection of incident sound from distributed inhomogeneities on the basilar membrane (i.e., cochlear roughness). In this model, the amplitude of the OAE is increased with increased height and breadth of the traveling wave. A more active newborn amplifier would produce enhanced gain, a taller and broader wave, and consequently an increased coherence of the backscattered wavelets. If the mechanisms of cochlear reflection are strongly functional at birth, other reflection emissions (CEOAEs, SFOAEs, and SOAEs) should also be more robust in infants. This, indeed, is a common finding (Norton and Widen 1990; Prievé 1992; Kalluri et al. 2011).

What process might provide stronger cochlear amplification in the neonatal cochlea? The idea of a functional overshoot period is plausible. The cochlear amplifier gain and tuning may be excessive during a brief developmental interlude, similar to what had been described in neonatal gerbils from 23 to 29 days after birth (Mills and Rubel 1996). The source of this overshoot in gerbils appears related to changes in the endocochlear potential (EP). Mills and Rubel hypothesized an adaptation mechanism that adjusts the “set-point” of the gain for optimal processing while reserving excess power. It is not known when the EP becomes adult-like in human infants because this requires invasive methodology. However, a transient overshoot period of similar origin, characterized by exuberant cochlear amplifier activity, is possible. A second possibility is that the pristine neonatal cochlea, as of yet untarnished by noise and ototoxins is at the pinnacle of performance, manifesting optimally functional cochlear amplification at birth, in contrast to an adult cochlea already impacted by natural aging, common ototoxins and noise. In this scenario, the newborn cochlea is *non-adultlike* but not *immature* since a subtle, subclinical degradation of the adult cochlear amplifier would be the source of age effects. Third, it is also possible that the efficiency of the medial efferent-OHC synapses may not be fully developed, and thus efferent-mediated modulation of OHC motility may not be effective, leading to somewhat unregulated gain and sharper tuning for a brief developmental period.

Another cluster of findings consistent with a strong neonatal cochlear amplifier is excessively narrow DPOAE suppression tuning observed in infants. Active processes are thought to underlie exquisite frequency tuning in the human cochlea. Recall that when the slope of suppression growth was examined in adults and infants for $f_2 = 6$ kHz, low-frequency side suppressors ($< f_2$) showed elevated suppression thresholds and compressive growth of suppression, consistent with narrower DPOAE suppression tuning (see Figs. 2.9 and 2.10). Because suppression growth is measured as a slope value derived re: each subject’s own suppression threshold, it is difficult to explain these results with simple attenuation of sound pressure level through an immature newborn middle ear. One possibility is that the elevated suppression thresholds for suppressors $< f_2$ (i.e., newborn functions shifted rightward on abscissa of Fig. 2.9; upper left panel) produced abbreviated growth functions, making it difficult to obtain a comprehensive picture of suppression growth.

Elevated suppression thresholds *could* be explained by lower levels reaching the infant cochlea, thereby leaving the middle ear as a possible contributor to these findings.

To scrutinize the impact of middle-ear development on DPOAE STCs, Abdala et al. (2007) examined correlations between measurements of DPOAE suppression tuning at $f_2 = 6$ kHz and selected acoustic transfer functions measured in the ear canal (admittance and energy reflectance at various frequencies). All measurements were performed longitudinally in the same group of infants from birth through 6 months of age. A relatively weak correspondence between cochlear and middle-ear indices was observed over the first half year of life. Of 75 correlations generated (5 DPOAE variables, 5 admittance variables, 3 frequencies), the admittance phase correlated most strongly with the tip-to-tail ratio of infant DPOAE STCs and could account for a maximum 25% of variance in DPOAE suppression tuning with age. Susceptance at 5.7 kHz could account for 18% of the variance with age. These weak associations might be due to relatively small numbers of infant subjects ($n = 20$) or noise intrinsic to each data set, but they do not suggest that changes in acoustic transfer functions measured in the ear canal can fully account for the developmental changes noted in DPOAE suppression tuning during the first months of life. As can be appreciated in any longitudinal study, a limitation was that the oldest infant age group was only 6 months with no other age groups tested between this age and adulthood.

Non-adultlike DPOAE suppression growth and narrower suppression tuning may be consistent with a vigorous cochlear amplifier in newborns. However, it is not likely that active processes underlying frequency selectivity, such as OHC somatic motility, remain immature into the sixth month of life. Morphological data do not support a prolonged postnatal time course for maturation of human OHC function. As speculated earlier, it is feasible that a transient overshoot period occurs as part of an EP-related adaptation mechanism, or that the age differences noted are explained by an optimally functional newborn sensory organ contrasted with a “well-worn” adult cochlea. As presented in Sect. 4.2, compelling arguments and models also exist that explain much of the infant DPOAE suppression tuning based solely on immaturity of ear-canal and middle-ear factors.

4.1.2 Cochlear Apex

Readers will recall that the slope of DPOAE phase was steeper in newborns than in adults for low-frequency signals coded in the apical half of the cochlea (see Fig. 2.6). Morphologically, the apex of the cochlea develops last and the apical region retains some immature-like (re: the base) features into adulthood (Pujol et al. 1998). In adult guinea pigs, among other small laboratory animals, the apex does not manifest fully nonlinear behavior or provide effective cochlear amplification (Cooper and Rhode 1995). In light of this relatively prolonged apical maturation, what might a prolonged DPOAE group delay and steeper slope of phase in the low frequencies indicate for newborn cochlear function? Because the middle ear impacts stimulus levels during forward transmission, it is first important to quantify the level

dependence of DPOAE phase, in particular for frequencies below 1.5 kHz where age effects have been observed. Toward this aim, DPOAE phase and individual component features were characterized in normal-hearing young adults at four stimulus levels (Abdala et al. 2011a). The DPOAE phase slope measured in the low-frequency segment (0.5–1.4 kHz) of phase-frequency functions did not show significant level dependence. In addition, once DPOAE components were separated, the distortion-component phase (which dominates the phase behavior shown in Fig. 2.6 and hence drives the phase trend in these data) showed no level dependence either. This result suggests that lower levels driving the infant cochlea (due to inefficient forward transmission through the newborn middle ear) do not easily explain steeper DPOAE phase observed in the apical half of the neonatal cochlea.

The fact that a steep phase gradient in newborns is observed in the apical region of the cochlea is intriguing and may be suggestive of immaturities in cochlear scaling. To understand cochlear scaling symmetry, it is necessary to refer back to tonotopic representation on the basilar membrane. Signals of different frequencies have to “travel” different distances along the cochlear partition to their characteristic frequency (CF) place. Scaling symmetry implies that the number of cycles to the CF place is relatively independent of frequency, and signals of all frequencies accumulate approximately the same amount of total phase (cycles) at their unique CF site. Because the cochlear frequency-place map is exponential, the traveling wave envelope is simply transposed along the basilar membrane for different CFs, producing shift similarity (Zweig 1976). The property of scaling can be observed and gauged with DPOAE phase. In a scaled system, when a fixed f_2/f_1 ratio is presented, the relative phases of the primary tones are constant throughout the frequency range; consequently DPOAE phase (which is calculated re: stimulus tone phase) is constant as well (Shera and Guinan 1999; Shera et al. 2000). This DPOAE phase invariance is thought to reflect local cochlear scaling symmetry.

In the mammalian auditory system, a breakdown in scaling in the apical half of the cochlea has been described (Zweig 1976; Shera et al. 2000, 2010). Such a break can be noted in the DPOAE data presented in lower panel of Fig. 2.6, where DPOAE phase transitions from flat in the mid- and high-frequencies to more rapidly cycling in the low-frequencies. Physiological evidence supports the hypothesis that the cochlear apex functions differently than the base (Cooper and Rhode 1995; Nowotny and Gummer 2006; Shera et al. 2010). The apical-basal transition derived from DPOAE phase versus frequency functions is centered between 1 and 1.4 kHz in adults and infants (depending on how it is measured) as shown in Fig. 2.6. Of interest here is that newborns exhibit a more marked break from scale invariance than adults suggesting a relatively prolonged course for the maturation of apical basilar membrane motion. The increased steepness of phase below this apical–basal transition frequency could imply a more marked deviation in the exponential relationship between frequency and place in the newborn apex. It might also indicate an age difference in the broadening of filters in the apical half of the cochlea. These interpretations, though intriguing, are speculative and warrant further consideration and experimentation.

DPOAE phase believed to mirror basilar membrane phase and motion, shows age differences for low-frequency signals, whereas cochlear amplification in the

infant cochlea appears to be mature at birth. The developmental sequence observed in most mammals is that passive motion of the basilar membrane develops before micromechanical aspects of cochlear function. Some relatively recent findings in mouse support this sequence (Song et al. 2008), though exceptions in gerbil have been reported (Overstreet et al. 2002). If DPOAE phase is determined by the material properties of the basilar membrane, does its immaturity in newborns suggest immature traveling wave properties in the apical half of the cochlea? There is a dearth of information about mass and stiffness characteristics of the basilar membrane in human fetal tissue. Therefore this issue cannot be adequately addressed with available literature.

Theories of cochlear development (and models of cochlear mechanics in general) have almost completely been formulated from observations in the mid- and high-frequency regions of the cochlea, yet the frequency range below 2 kHz is important to human communication and includes frequencies salient to speech intelligibility. Recent reports of DPOAE phase differences between newborns and adults at low frequencies suggest that studying the apical half of the human cochlea around the time of birth may prove fruitful in elucidating immaturities in cochlear mechanics. The difficult task will be to develop paradigms that effectively target and assay low-frequency features of cochlear function in the challenging newborn population.

4.2 *Ear-Canal and Middle-Ear Explanations*

Possible sources of OAE immaturity may be better understood through considering how immaturities in ear-canal and middle-ear functioning affect OAEs. Elements of a model were introduced in Sects. 2.2.4–2.2.6, in which the goal for a general hearing experiment was to describe forward transmission of sound from the canal to the cochlea. Measurements of the forward ear-canal transfer function level L_{FE} and reverse ear-canal transfer function level L_{RE} were described. A total forward transfer function level L_F was defined as the sum of L_{FE} and a forward middle-ear transfer function level L_{FM} . The following description of bidirectional OAE transmission also introduces a total reverse transfer function level L_R as the sum of L_{RE} and a reverse middle-ear transfer function level L_{RM} .

4.2.1 **Modeling DPOAE Response Growth**

As described in Sect. 3, the largest differences in DPOAE response growth between measurements in young adult and infant ears occurred at a stimulus frequency f_2 of 6 kHz. This response growth is shown in Fig. 2.7 as a DPOAE I/O function, in which f_2/f_1 was fixed at 1.2, and L2 was always 10 dB below L1. Here, a model is adopted that cochlear mechanics are mature at birth, so that all age differences in DPOAE I/O functions can be accounted for by ear-canal and middle-ear immaturities. The goal was to describe maturational changes in these DPOAE data at 6 kHz as a function of maturational changes in $L_F, L_{FM}, L_R,$ and L_{RM} , in which maturational

changes in L_{FE} and L_{RE} were independently calculated (see Fig. 2.3). The model calculated relative levels ΔL_F , ΔL_{FM} , ΔL_R , and ΔL_{RM} , respectively, in which the each relative level was defined as the level difference in an infant group minus the adult group. For example, ΔL_F was the difference in L_F for that infant group minus L_F for the adult group. Even though L_F and other transfer functions could not be directly measured for any age group (because an invasive measurement would be needed; see Sect. 2.2.6), it was possible to calculate the relative total forward transfer function level ΔL_F based on DPOAE response growth data.

In this approach, horizontal translation of a DPOAE I/O function toward higher stimulus levels corresponded to reduced forward transmission of the stimuli eliciting the DPOAE. Vertical translation of the I/O function toward higher DPOAE levels corresponded to increased reverse transmission of the DPOAE. The amounts of horizontal and vertical translations needed to best align the I/O function at each infant age group (full term, 3, 4, 5, and 6 months) with the “reference” adult I/O function provided ΔL_F at the f_1 stimulus frequency and ΔL_R at the DPOAE frequency ($2f_1 - f_2$) (Abdala and Keefe 2006). The fact that adequate fits were obtained for each infant age group confirmed the hypothesis of cochlear maturity.

Maturation differences were then calculated using data in Fig. 2.3. The relative forward ear-canal transfer function level ΔL_{FE} was calculated from 0.25 to 8 kHz as the difference in L_{FE} for each infant group relative to adults; the relative reverse ear-canal transfer function level ΔL_{RE} was similarly calculated. Only ΔL_{FE} at $f_1 = 6$ kHz and ΔL_{RE} at the DP frequency were used in the model.

Inasmuch as $L_F = L_{FE} + L_{FM}$, then $\Delta L_F = \Delta L_{FE} + \Delta L_{FM}$, so that the relative forward middle-ear transfer function level was calculated by $\Delta L_{FM} = \Delta L_F - \Delta L_{FE}$. These relative forward levels are plotted in the top panel of Fig. 2.11 for each age group according to the mean level and ± 1 standard error (SE) of the mean. The ear-canal component ΔL_{FE} was close to 0 dB for term infants but increased with increasing age. This is a clear sign that immaturities related to ear-canal function and to the absorption and reflection properties of the tympanic membrane have not diminished by age 6 months. Conversely, the middle-ear component ΔL_{FM} showed 16 dB of attenuation in the term infant that was reduced to 8 dB of attenuation by age 6 months. Almost accidentally, the sum of these two processes represented by ΔL_F is within ± 1 SE of 0 dB, which would appear to suggest maturity in forward transmission at 6 kHz in the absence of this decomposition. In fact, ΔL_{FE} varied rapidly with frequency between 2 and 8 kHz (because of the rapid fluctuation of L_{FE} for the adult ear response in the top panel of Fig. 2.3), so that the near cancellation at 6 kHz would be unlikely to occur at other frequencies.

By an analogous argument, the relative reverse middle-ear transfer function levels were calculated by $\Delta L_{RM} = \Delta L_R - \Delta L_{RE}$. Each is plotted (mean and SE) in the bottom panel of Fig. 2.11. In contrast to forward transmission, each relative reverse level was largest and positive for term infants, and each tended to decrease with increasing age out to 6 months. A positive ΔL_R signifies that OAE levels were boosted in infants compared to adults. Most of this effect comes from an increase in ΔL_{RM} , because, as is evident from comparing infant and adult values of L_{RE} in the bottom

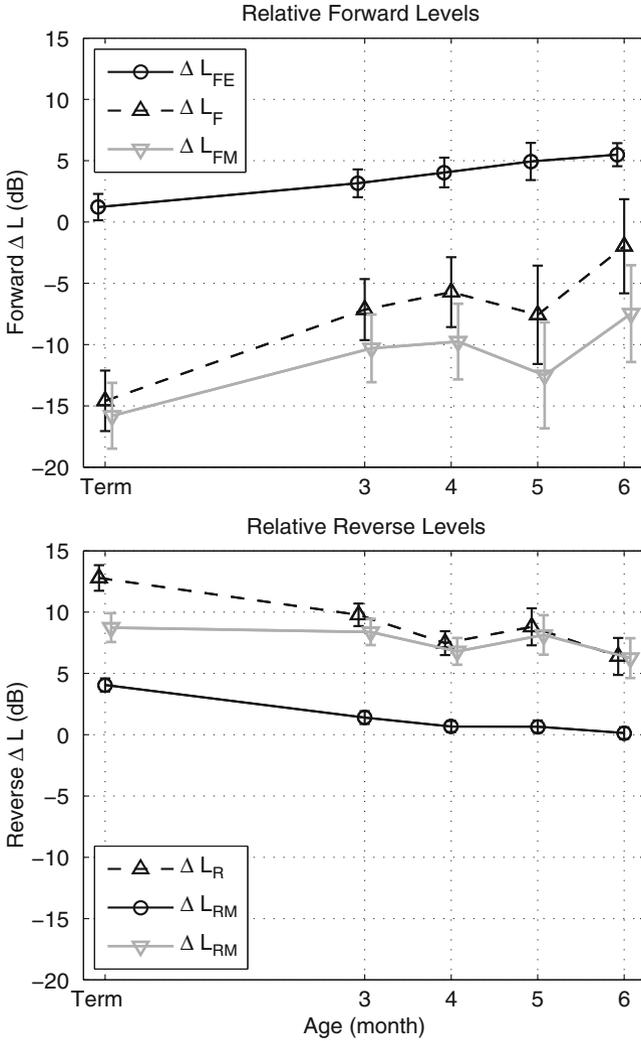


Fig. 2.11 Relative ear-canal and middle-ear transfer-function levels corresponding to DPOAEs at $f_2=6$ kHz as functions of infant age. *Top*: Relative forward transfer-function levels. *Bottom*: Relative reverse transfer-function levels [Parts of the figure were originally published in Keefe and Abdala (2007)]

panel of Fig. 2.3, ΔL_{RE} remained close to 0 dB. ΔL_{RM} was immature in all age groups, decreasing from 9 dB in term infants to 7 dB at age 6 months (Fig. 2.11).

An interesting finding was that the infant middle ear attenuated the stimulus in the forward direction ($\Delta L_{FM} < 0$) but boosted the DPOAE in the reverse direction ($\Delta L_{RM} > 0$). This appears counterintuitive inasmuch as the middle ear might be expected to have an attenuation in both directions or neither. This sign difference is explained in the model by its prediction that ΔL_{RM} is related to ΔL_{FM} via the relative

area level (see Fig. 2.2), that is, $\Delta L_{RM} = \Delta L_{FM} - \Delta L_a$. ΔL_a was always positive and as large as 17 dB in full-term infants, which is the factor producing an overall positive ΔL_{RM} . The small ear-canal area in infants accounts for this boost in reverse transmission. From the perspective of a general hearing experiment (i.e., not concerned with reverse transmission), the relevant finding is that forward transmission was attenuated in the infant ear compared to the adult ear (at least at 6 kHz).

Immaturities in ear-canal and middle-ear function would also affect other OAE types. For example, it would be possible model immaturities in CEOAE and SFOAE responses using a similar approach to that used for DPOAEs. CEOAE levels are larger in infants than in adults, consistent with reverse-transmission effects described above for DPOAEs. However, estimating the relative forward and reverse levels (i.e., ΔL_F and ΔL_R) from a CEOAE I/O function is more difficult. DPOAE I/O functions are nonmonotonic so that it was possible to translate an infant DPOAE I/O function to best match an adult DPOAE I/O function. CEOAE I/O functions are much more nearly monotonic functions of stimulus level, which would complicate the procedure used for DPOAEs.

4.2.2 DPOAE Suppression

A measurement of a DPOAE I/O function implicates ear-canal and middle-ear functioning via forward transmission of stimulus energy at f_1 and f_2 , and reverse transmission at the DP frequency. Measuring the suppression of a DPOAE further implicates ear-canal and middle-ear functioning via forward transmission level differences of each suppressor tone at its suppressor frequency. Given the ability to explain immaturities in DPOAE I/O functions at $f_2 = 6$ kHz in terms of immaturities in the conductive pathway, several studies (Abdala and Keefe 2006; Abdala et al. 2007; Keefe and Abdala 2007) explored whether conductive immaturities also explain immaturities in DPOAE suppression tuning.

As described in Sect.3.6 and shown in Figs. 2.8 and 2.10, DPOAE STCs in infants were most different from STCs in adults at $f_2 = 6$ kHz. For example, the tip-to-tail level reported by Abdala and Keefe (2006) was 27 dB in infants compared to only 15 dB in adults at reference L1–L2 levels of 65–55 dB SPL. As shown in the top panel of Fig. 2.11, the attenuation in forward transmission level ($-\Delta L_F$) was 15 dB in full-term neonates. DPOAE STCs in infants were compared to STCs in adults at probe levels that were reduced by 15 dB, that is, at L1–L2 of 50–40 dB SPL, to calibrate equal cochlear levels based on ΔL_F . The DPOAE STC in adults at this reduced probe level had a tip-to-tail level of 26 dB, nearly equal to that for the infant STC at the reference level.

With immaturities in forward transmission of suppressor tones accounted for, it remained to correct for immaturities in the magnitude of the DPOAE that was being suppressed. The tip of the adult STC at the reduced probe level was about 15 dB below the tip of the infant STC at the reference probe level. The “best fit” in the tip region was obtained by adding 15 dB to the adult STC. This amount was within a

couple dB of the boost in reverse transmission of $\Delta L_R = 13$ in full-term infants (bottom panel of Fig. 2.11). Clearly, at $f_2 = 6$ kHz, immaturities in ear-canal and middle-ear function explained much of the difference in the DPOAE STCs of term infants and adults. This result is consistent with the theory that cochlear function is mature, and ear-canal and middle-ear functions are immature, in the human neonate.

As noted in Sect. 4.1, several frequency-specific components of admittance and reflectance did not correlate strongly with DPOAE STC indices in the first 6 months of life (Abdala et al. 2007). This is not surprising inasmuch as these components did not directly measure the total forward or reverse transmission through the ear canal and middle ear (such a measurement would require invasive techniques as explained in Sect. 2.2). Each component is a measure of conductive function as viewed in a mid-canal location that may not highly correlate with transmission to or from the cochlea.

Even though DPOAE suppression in infants would be influenced by forward transmission at each suppressor frequency, a single measurement of forward transmission at the probe frequency (f_2) would not specify the transmission over the range of suppressor frequencies. Suppressor levels are assessed using pressure measurements at the probe, but acoustic standing-wave distributions of ear-canal pressure differ substantially between infants and adults over the frequency range of the DPOAE STC. It would be preferable to measure STCs using a variable that is not influenced by standing waves in the ear canal.

As shown in Fig. 2.11, even though forward transmission (ΔL_F) at 6 kHz was estimated to be nearly adult-like in 6-month-olds, the low-frequency flank of the mean DPOAE STC in Fig. 2.10 remained immature in 6-month-olds. This immaturity has been cited as possible evidence for residual immaturities in cochlear mechanics (Dhar and Abdala 2007; Abdala and Dhar 2010). Maturation differences in forward transmission at any suppressor frequency might produce differences in the resulting DPOAE STC, but data at the suppressor frequencies were not included in the model that predicted ΔL_F (DPOAE I/O function data would be required at each suppressor frequency). One problem is the presence of age-related differences in ear-canal standing waves at each suppressor frequency, so that the use of sound pressure level can be misleading. It is possible to control partially for these differences by measuring the DPOAE STC using a different acoustic ear-canal variable—absorbed sound power rather than sound pressure.

Keefe and Schairer (2011) described differences between measuring a SFOAE STC as a function of absorbed sound power level compared to SPL. The shape of the SFOAE STC recorded in adult ears at a stimulus frequency of 8 kHz was significantly changed when absorbed power level rather than SPL was used to specify the level of the stimuli generating and suppressing the SFOAE. The underlying causes were the presence of acoustic standing waves in the adult ear canal that produced maximal effects around 4 kHz, and the increased efficiency of absorbing sound energy by the middle ear at frequencies between 2 and 4 kHz compared to lower and higher frequencies up to 8 kHz. The effects of these standing waves in adults are shown for L_{FE} near 4 kHz (see top panel, Fig. 2.3).

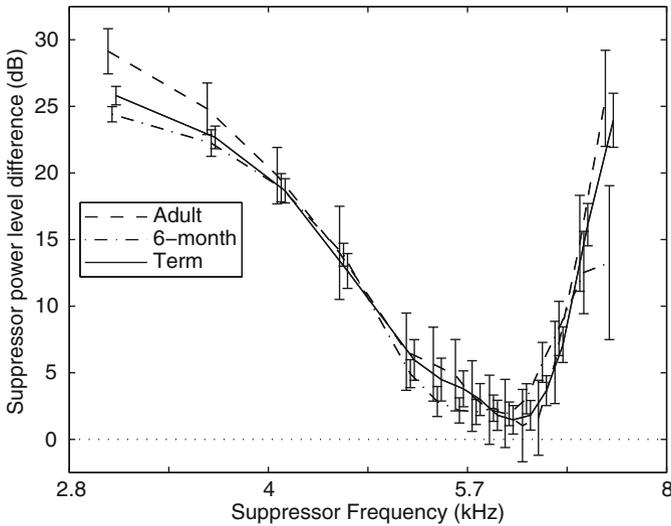


Fig. 2.12 The suppressor power level difference is the difference between the absorbed-power DPOAE suppression tuning curve (at $f_2=6$ kHz) producing the criterion decrement of the DPOAE and the stimulus absorbed power level averaged over f_1 and f_2 . The suppression tuning curves are based on DPOAEs recorded at the same set of stimulus SPLs for each f_1 and f_2 . The mean suppressor power level difference is plotted for groups of adults, 6-month-olds, and newborns as a function of suppressor frequency. The error bars represent ± 1 standard error of the absorbed-power DPOAE suppression tuning curve [This figure was originally published in Keefe and Abdala (2011)]

Keefe and Abdala (2011) described maturational differences between DPOAE STC measurements as a function of absorbed sound power level compared to measurements based on SPL. These were based on DPOAEs recorded at the same set of stimulus SPLs in full-term infants, infants at age 6 months and adults. The DPOAE STC was first transformed from a criterion SPL at each suppressor frequency to its absorbed power level. Then, it was translated based on the conversion from the stimulus SPLs to the stimulus absorbed power level averaged over the stimulus frequencies f_1 and f_2 . This STC conversion controlled for maturational differences in the unsuppressed DPOAE generated using fixed stimulus levels across the age groups (stimulus frequencies and levels were the same as for the STC results shown in Fig. 2.10). The resulting suppressor power level difference was defined as the difference between the absorbed-power DPOAE STC (at $f_2=6$ kHz) producing the criterion decrement of the DPOAE, and this stimulus absorbed power level.

The mean and SE of the suppressor power level differences are plotted in Fig. 2.12. DPOAE STCs were similar across age at nearly all suppressor frequencies, once the SPLs of the stimulus and suppressor tones were converted to absorbed power level (i.e., compare Figs. 2.10 and 2.12). Normalizing the mean suppressor power level difference using the stimulus absorbed power level matched the STC tip levels across age in Fig. 2.12. These results suggest that the cochlear mechanics that underlie DPOAE suppression are substantially mature in full-term infants.

Thus, it appears possible to explain the differences between infants and adults in DPOAE STCs at a probe frequency of 6 kHz in terms of ear-canal and middle-ear immaturities. A particularly important factor is that the L_{FE} has a strong frequency resonance in adults in the octave range below 6 kHz, whereas it is slowly varying in infants (top panel, Fig. 2.3). This is an instructive example of why it might be helpful in future research involving measurements in infant and adult ears to calibrate ear-canal stimuli in terms of absorbed power level rather than SPL. This would also apply in experiments measuring electrophysiological or behavioral responses.

4.2.3 Immaturities in SOAEs

Age differences in prevalence, level, and spectra of SOAEs may also be explained in terms of immaturities in ear-canal and middle-ear function. In terms of the standing wave theory of SOAE generation (see Sect. 4.1.1), SOAEs are influenced in a similar manner by the relative forward and reverse ear-canal transfer function levels L_{FE} and L_{RE} in infants relative to adults. SOAE age differences may be consistent with increased reverse transmission of energy from the cochlea to the ear canal in infants. Increased reverse transmission would boost the SOAE level, thus increasing the likelihood of detection for a fixed noise level. This increased SOAE level would correctly predict an increased prevalence of SOAEs in infants relative to adults. While no forward-transmitted stimulus elicits the SOAE, multiple internal reflections of the SOAE would create a cascade of forward-transmitted signals. Similarly, SOAEs are also influenced by the relative forward and reverse middle-ear transfer-function levels ΔL_{FM} and ΔL_{RM} . As with other OAE types, SOAE levels are boosted by the smaller ear-canal area and shorter length in infants relative to adults.

4.2.4 Discussion

The findings that immaturities in ear-canal and middle-ear function are sufficient to adequately explain immaturities in DPOAE response growth and DPOAE suppression suggest that they may help explain immaturities in other types of DPOAE responses. The underlying idea as regards forward transmission is that changes in cochlear input level, which are produced by functional immaturities in more peripheral parts of the ear, lead to changes in cochlear output level resulting from the compressively nonlinear response on the basilar membrane. More peripheral immaturities in reverse transmission filter the OAE signal measured in the ear canal and change the strength of any multiple internal reflections.

Whether residual immaturities in the cochlear structures of infants lead to detectable functional differences in OAE responses is a challenging and interesting scientific problem. It will be important in such future experiments to consider the

impact of immaturities in ear-canal and middle-ear function that will otherwise tend to dominate the immaturities observed in OAE responses.

5 Summary

The maturational processes underlying the acoustical and mechanical functioning of the human ear and the structures serving those functions were described in the first part of this chapter. The development of ear-canal and middle-ear structures is not yet complete in the full-term infant, but continues at least up to the onset of puberty. While postnatal growth of these peripheral structures is rapid during the first year of life, more study is needed concerning the postnatal maturation of soft tissue such as the tympanic membrane. The substantial postnatal maturation of the ear canal and middle ear produces functional immaturities in how the ear canal and middle ear receive, absorb, and transmit sound energy to the cochlea. A description of the acoustical–mechanical coupling of power flow through the ear can account for maturational differences in how sound presented in the ear canal is transmitted to the inner ear, based on a detailed examination of experimental data collected in adult ears and in the ears of infants throughout the age span. Understanding the maturational processes in terms of power flow through the peripheral auditory system is important to understanding maturation processes in hearing and in interpreting measurements of physiological responses such as OAEs and auditory brain stem responses.

The second part of this chapter reviewed maturation of cochlear structure and function as explained by measures of otoacoustic emissions. In contrast to outer- and middle-ear anatomy, cochlear structures appear to be substantially completed sometime in the third trimester but the final stages are poorly delineated and little is known about changes in the physical properties of the basilar membrane during the early postnatal period. This section provided a careful examination and prudent interpretation of OAEs recorded in human infants to better understand underlying cochlear mechanics during maturation. All indications from these findings suggest the cochlea of newborns produces robust cochlear amplification, yet age differences were identified in the relative magnitude of DPOAE components and in the DPOAE phase gradient, particularly in the apical half of the newborn cochlea. These results warrant further consideration and study. It is clear from the information presented in this chapter that any such endeavor must take into account the relatively long time course for the maturation of outer- and middle-ear properties and to consider how these factors impact OAE recordings from infants and children.

Though it would have been convenient (for readers, perhaps) to outline a more definitive timeline for the development of human cochlear function and provide more firm interpretations, the complexity of this maturational process and the influence of noncochlear factors such as the conductive system and the medial efferent system make this a formidable task. OAE-based measures provide a glimpse into the human cochlea during the earliest segments of perinatal and postnatal life though

it is clear that they provide this glimpse through a two-way mirror of ear-canal and middle-ear function.

Acknowledgments This work was supported by a grant from the National Institutes of Health, R01 DC003552 (CA) and by the House Research Institute. C. Abdala thanks Dr. Rangasamy Ramanathan, Chief of Neonatology at the University of Southern California, Keck School of Medicine for continued support of neonatal auditory research and Dr. Silvia Batezati for assistance with the preparation of this chapter.

References

- Abbas, P. J., & Sachs, M. B. (1976). Two-tone suppression in auditory nerve fibers: Extension of a stimulus-response relationship. *Journal of the Acoustical Society of America*, 59, 112–122.
- Abdala, C. (1996). Distortion product otoacoustic emission ($2f_1-f_2$) amplitude as a function of f_2/f_1 frequency ratio and primary tone level separation in human adults and neonates. *Journal of the Acoustical Society of America*, 100, 3726–3740.
- Abdala, C., Sininger, Y. S., Ekelid, M., & Zeng, F.-G. (1996). Distortion product otoacoustic emissions suppression tuning curves in human adults and neonates. *Hearing Research*, 98, 38–53.
- Abdala, C. (1998). A developmental study of distortion product otoacoustic emission ($2f_1-f_2$) suppression in humans. *Hearing Research*, 121, 125–138.
- Abdala, C. (2000). Distortion product otoacoustic emission ($2f_1-f_2$) amplitude growth in human adults and neonates. *Journal of the Acoustical Society of America*, 107, 446–456.
- Abdala, C. (2001). Maturation of the human cochlear amplifier: Distortion product otoacoustic emission suppression tuning curves recorded at low and high primary tone levels. *Journal of the Acoustical Society of America*, 110, 1465–1476.
- Abdala, C. (2003). A longitudinal study of distortion product otoacoustic emission ipsilateral suppression and input/output characteristics in human neonates. *Journal of the Acoustical Society of America*, 114, 3239–3250.
- Abdala, C., & Chatterjee, M. (2003). Maturation of cochlear nonlinearity as measured by distortion product otoacoustic emission suppression growth in humans. *Journal of the Acoustical Society of America*, 114, 932–943.
- Abdala, C., & Dhar, S. (2010). Distortion product otoacoustic emission (DPOAE) phase and component analysis in human newborns. *Journal of the Acoustical Society of America*, 127(1), 316–325.
- Abdala, C., & Keefe, D. H. (2006). Effects of middle-ear immaturity on distortion product otoacoustic emission suppression tuning in infant ears. *Journal of the Acoustical Society of America*, 120, 3832–3842.
- Abdala, C., Ma, E., & Sininger, Y. (1999). Maturation of medial efferent system function in humans. *Journal of the Acoustical Society of America*, 105, 2392–2402.
- Abdala, C., Keefe, D. H., & Oba, S. I. (2007). Distortion product otoacoustic emission suppression tuning and acoustic admittance in human infants: Birth through six months. *Journal of the Acoustical Society of America*, 121, 3617–3627.
- Abdala, C., Oba, S. I., & Ramanathan, R. (2008). Changes in the DP-gram during the preterm and early-postnatal period. *Ear and Hearing*, 29, 512–523.
- Abdala, C., Mishra, S. R., & Williams, T. L. (2009). Considering distortion product otoacoustic emission fine structure in measurements of the medial olivocochlear reflex. *Journal of the Acoustical Society of America*, 125, 1584–1594.
- Abdala, C., Dhar, S., & Kalluri, R. (2011a). Level dependence of DPOAE phase is attributed to component mixing. *Journal of the Acoustical Society of America*, 129, 3123–3134.
- Abdala, C., Dhar, S., & Mishra, S. (2011b). The breaking of cochlear scaling symmetry in human newborns and adults. *Journal of the Acoustical Society of America*, 129, 3104–3115.

- Abdala, C., Mishra, S. R., Batezati, S. C., & Wiley, J. M. (2011c). Maturation of the MOC reflex in humans: 12 years later. Abstract 376. *Association for Research in Otolaryngology* 34th Midwinter Meeting, February 19–23, 2011, Baltimore, MD.
- Anson, B. J., & Donaldson, J. A. (1981). *Surgical anatomy of the temporal bone and ear*. Philadelphia: W. B. Saunders.
- Birnholz, J. C., & Benacerraf, B. R. (1983). The development of human fetal hearing. *Science*, 222, 516–518.
- Bonfils, P., Uziel, A., & Narcy, P. (1989). The properties of spontaneous and evoked acoustic emissions in neonates and children: A preliminary report. *Archives of Oto-rhino-laryngology*, 246, 249–251.
- Bonfils, P., Avan, P., Francois, M., Trotoix, J., & Narcy, P. (1992). Distortion-product otoacoustic emissions in neonates: Normative data. *Acta Oto-Laryngologica*, 112, 739–744.
- Bredberg, G. (1968). Cellular pattern and nerve supply of the human organ of Corti. *Acta Oto-Laryngologica Supplementum*, 236, 1.
- Briennesse, P., Antenius, L. J., Maertzdorf, W. J., Blanco, C. E., & Manni, J. J. (1997). Frequency shift of individual spontaneous otoacoustic emissions in preterm infants. *Pediatric Research*, 42, 478–483.
- Brown, A. M., & Kemp, D. T. (1984). Suppressibility of the $2f_1-f_2$ stimulated acoustic emissions in gerbil and man. *Hearing Research*, 13, 29–37.
- Brown, A. M., Sheppard, S., & Russell, P. (1994). Acoustic distortion products (ADP) from the ears of term infants and young adults using low stimulus levels. *British Journal of Audiology*, 28, 273–280.
- Brown, A. M., Sheppard, S., & Russell, P. (1995). Differences between neonate and adult cochlear mechanical responses. *Auditory Neuroscience*, 1, 169–181.
- Burns, E. M., Arehart, K. H., & Campbell, S. L. (1992). Prevalence of spontaneous otoacoustic emissions in neonates. *Journal of the Acoustical Society of America*, 91, 1571–1575.
- Burns, E. M., Campbell, S. L., & Arehart, K. H. (1994). Longitudinal measurements of spontaneous otoacoustic emissions in infants. *Journal of the Acoustical Society of America*, 95, 385–394.
- Collet, L. (1993). Use of otoacoustic emissions to explore the medial olivocochlear system in humans. *British Journal of Audiology*, 27, 155–159.
- Cooper, N., & Rhode, W. (1995). Nonlinear mechanisms at the apex of the guinea pig cochlea. *Hearing Research*, 82, 225–243.
- Crelin, E. S. (1973). *Functional anatomy of the newborn*. New Haven and London: Yale University Press.
- Deeter, R., Abel, R., Calandrucchio, L., & Dhar, S. (2009). Contralateral acoustic stimulation alters the magnitude and phase of distortion product otoacoustic emissions. *Journal of the Acoustical Society of America*, 126, 2413–2424.
- Delgutte, B. (1990). Two-tone suppression in auditory nerve fibers: Dependence on suppressor frequency and level. *Hearing Research*, 49, 225–246.
- Dempster, J. H., & Mackenzie, K. (1990). The resonance frequency of the external auditory canal in children. *Ear and Hearing*, 11, 296–298.
- Dhar, S., & Abdala, C. (2007). A comparative study of DPOAE fine structure in human newborns and adults with normal hearing. *Journal of the Acoustical Society of America*, 122, 2191–2202.
- Eby, T. L., & Nadol, J. B., Jr. (1986). Postnatal growth of the human temporal bone: Implications for cochlear implants in children. *The Annals of Otolaryngology, Rhinology, and Laryngology*, 95, 356–364.
- Eggermont, J., Brown, D., Ponton, C., & Kimberley, B. (1996). Comparisons of DPE and ABR traveling wave delay measurements suggest frequency specific synapse maturation. *Ear and Hearing*, 17, 386–394.
- Feeney, M. P., & Keefe, D. H. (2011). Physiological mechanisms assessed by aural acoustic transfer functions. In K. Tremblay & R. Burkard (Eds.), *Translational perspectives in auditory neuroscience* (in press). San Diego: Plural.

- Fex, J. H. (1962). Augmentation of cochlear microphonic by stimulation of efferent fibers to the cochlea. *Acta Oto-Laryngologica*, 50, 540–541.
- Galambos, R., & Hecox, K. E. (1978). Clinical applications of the auditory brain stem response. *Otolaryngologic Clinics of North America*, 11, 709–722.
- Guinan, J. J., Jr. (2006). Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and Hearing*, 27, 589–607.
- Holte, L., Margolis, R. L., & Cavanaugh, R. M., Jr. (1991). Developmental changes in multifrequency tympanograms. *Audiology*, 30, 1–24.
- Igarishi, Y. (1980). Cochlea of the human fetus: A scanning electron microscope study. *Archivum Histologicum Japonicum*, 43, 195–209.
- Ikui, A., Sando, I., Sudo, M., & Fujita, S. (1997). Postnatal change in angle between the tympanic annulus and surrounding structures. *The Annals of Otology, Rhinology and Laryngology*, 106, 33–36.
- Ikui, A., Sando, I., Haginomori, S., & Sudo, M. (2000). Postnatal development of the tympanic cavity: A computer-aided reconstruction and measurement study. *Acta Oto-Laryngologica*, 120, 375–379.
- Kalluri, R., Abdala, C., Mishra, S., & Gharibian, L. (2011). Stimulus-frequency otoacoustic emissions in human newborns. Abstract 369. *Association for Research in Otolaryngology* 34th Midwinter Meeting, February 19–23, 2011, Baltimore, MD.
- Keefe, D. H., & Abdala, C. (2007). Theory of forward and reverse middle-ear transmission applied to otoacoustic emissions in infant and adult ears. *Journal of the Acoustical Society of America*, 121, 978–993.
- Keefe, D. H., & Abdala, C. (2011). Distortion-product otoacoustic-emission suppression tuning in human infants and adults using absorbed sound power. *Journal of the Acoustical Society of America – Express Letters*, 129, 108–113.
- Keefe, D. H., & Schairer, K. S. (2011). Specification of absorbed-sound power in the ear canal. *Journal of the Acoustical Society of America*, 129, 779–791.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in humans infants and adults. *Journal of the Acoustical Society of America*, 94, 2617–2638.
- Keefe, D. H., Bulen, J. C., Campbell, S. L., & Burns, E. M. (1994). Pressure transfer function and absorption cross section from the diffuse field to the human infant ear canal. *Journal of the Acoustical Society of America*, 95, 355–371.
- Keefe, D. H., Fitzpatrick, D., Liu, Y. W., Sanford, C. A., & Gorga, M. P. (2010). Wideband acoustic-reflex test in a test battery to predict middle-ear dysfunction. *Hearing Research*, 263, 52–65.
- Kemp, D. T. (1979). Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. *Archives of Oto-rhino-laryngology*, 224, 37–45.
- Knight, R. D., & Kemp, D. T. (2001). Wave and place fixed DPOAE maps of the human ear. *Journal of the Acoustical Society of America*, 109, 1513–1525.
- Kuypers, L. C., Decraemer, S. F., & Dirckx, J. J. J. (2006). Thickness distribution of fresh and preserved human eardrums measured with confocal microscopy. *Otology & Neurotology*, 27, 256–264.
- Lasky, R. E. (1998). Distortion product otoacoustic emissions in human newborns and adults. I. Frequency effects. *Journal of the Acoustical Society of America*, 103, 981–991.
- Lasky, R. E., Perlman, J., & Hecox, K. (1992). Distortion-product otoacoustic emissions in human newborns and adults. *Ear and Hearing*, 13, 430–441.
- Lavigne-Rebillard, M., & Pujol, R. (1986). Development of the auditory hair cell surface in human fetuses. A scanning electron microscopy study. *Anatomy and Embryology*, 174, 369–377.
- Lavigne-Rebillard, M., & Pujol, R. (1987). Surface aspects of the developing human organ of Corti. *Acta Oto-Laryngologica Supplementum*, 436, 43–50.
- Lavigne-Rebillard, M., & Pujol, R. (1990). Auditory hair cells in human fetuses: Synaptogenesis and ciliogenesis. *Journal of Electron Microscopy Technique*, 15, 115–122.
- Lim, D. J. (1970). Human tympanic membrane. An ultrastructural observation. *Acta Oto-Laryngologica* 70, 176–186.

- Lim, D. J., & Rudea, J. (1992). Structural development of the cochlea. In R. Romand (Ed.), *Development of the auditory and vestibular systems*, Vol. 2 (pp. 33–58). Amsterdam: Elsevier.
- Lonsbury-Martin, B., Harris, F., Hawkins, M., Stagner, B., & Martin, G. (1990). Distortion product otoacoustic emissions in humans: I. Basic properties in normal-hearing subjects. *Annals of Otology, Rhinology & Laryngology Supplementum*, 147, 3–14.
- Maison, S. F., & Liberman, M. C. (2000). Predicting vulnerability to acoustic injury with a noninvasive assay of olivocochlear reflex strength. *The Journal of Neuroscience*, 20, 4701–4707.
- Martin, G., Lonsbury-Martin, B., Prost, R., Scheinin, S., & Coats, A. (1987). Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contours and pure-tone exposures. *Hearing Research*, 28, 191–208.
- McLellan, M. S., & Webb, C. H. (1957). Ear studies in the newborn infant. *Journal of Pediatrics*, 51, 672–677.
- Micheyl, C., & Collet, L. (1996). Involvement of the olivocochlear bundle in the detection of tones in noise. *Journal of the Acoustical Society of America*, 99, 1604–1610.
- Mills, D. M., & Rubel, E. W. (1996). Development of the cochlear amplifier. *Journal of the Acoustical Society of America*, 100, 428–441.
- Morales B., Choi S. Y., & Kirkwood, A. (2002). Dark rearing alters the development of GABAergic transmission in visual cortex. *Journal of Neuroscience*, 22, 8084–8090.
- Morlet, T., Collet, L., Salle, B., & Morgon, A. (1993). Functional maturation of cochlear active mechanisms and of the medial olivocochlear system in humans. *Acta Oto-Laryngologica*, 113, 271–277.
- Morlet, T., Lapillonne, A., Ferber, C., Duclaux, R., Sann, L., Putet, G., Salle, B., & Collet, L. (1995). Spontaneous otoacoustic emissions in preterm neonates: Prevalence and gender effects. *Hearing Research*, 90, 44–54.
- Mountain, D. (1980). Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanisms. *Science*, 210, 71–72.
- Northern, J. L., & Downs, M. P. (1984). *Hearing in children*. Baltimore: Williams & Wilkins.
- Norton, S. J., & Widen, J. (1990). Evoked otoacoustic emissions in normal-hearing infants and children: emerging data and issues. *Ear and Hearing*, 11, 121–127.
- Nowotny, M., & Gummer, A.W. (2006). Electromechanical transduction: Influence of the outer hair cells on the motion of the organ of Corti. *HNO*, 54, 536–543.
- Okabe, K., Tanaka, S., Hamada, H., Miura, T., & Funai, H. (1988). Acoustic impedance measurement on normal ears of children. *Journal of Acoustical Society of Japan*, 9, 287–294.
- Overstreet, E. H. III, Temchin, A. N., & Ruggero, M. A. (2002). Passive basilar membrane vibrations in gerbil neonates: Mechanical bases of cochlear maturation. *Journal of Physiology*, 545, 279–288.
- Popelka, G. R., Karson, R. K., & Arjmand, E. M. (1995). Growth of the $2f_1-f_2$ distortion product otoacoustic emission for low-level stimuli in human neonates. *Ear and Hearing*, 16, 159–165.
- Prieve, B. A. (1992). Otoacoustic emissions in infants and children: Basic characteristics and clinical application. *Seminars in Hearing*, 13, 37–52.
- Prieve, B. A., Fitzgerald, T. S., & Schulte, L. E. (1997a). Basic characteristics of click-evoked otoacoustic emissions in infants and children. *Journal of the Acoustical Society of America*, 102, 2860–2870.
- Prieve, B. A., Fitzgerald, T. S., Schulte, L. E., & Kemp, D. T. (1997b). Basic characteristics of distortion product otoacoustic emissions in infants and children. *Journal of the Acoustical Society of America*, 102, 2871–2879.
- Puel, J. L., & Rebillard, G. (1990). Effect of contralateral sound stimulation on the distortion product $2f_1-f_2$: Evidence that the medial efferent system is involved. *Journal of the Acoustical Society of America*, 87, 1630–1635.
- Pujol, R., & Hilding, D. (1973). Anatomy and physiology of the onset of auditory function. *Acta Oto-Laryngologica*, 76, 1–11.
- Pujol, R., & Lavigne-Rebillard, M. (1985). Early stages of innervation and sensory cell differentiation in the human fetal organ of Corti. *Acta Oto-Laryngologica Supplementum*, 423, 43–50.

- Pujol, R., Carlier, E., & Lenoir, M. (1980). Ontogenetic approach to inner and outer hair cells functions. *Hearing Research*, 2, 423–430.
- Pujol, R., Zajic, G., Dulon, D., Raphael, Y., Altschuler, R. A., & Schacht, J. (1991). First appearance and development of motile properties in outer hair cells isolated from guinea-pig cochlea. *Hearing Research*, 57, 129–141.
- Pujol, R., Lenoir, M., Ladrech, S., Tribillac, F., & Rebillard, G. (1992). Correlation between the length of outer hair cells and the frequency coding of the cochlea. In Y. Cazals, L. Demany, & K. C. Horner (Eds.), *Auditory physiology and perception* (pp. 45–52). Oxford: Pergamon Press.
- Pujol, R., Lavigne-Rebillard, M., & Lenoir, M. (1998). Development of sensory and neural structures in the mammalian cochlea. In E. Rubel, A. Popper, & R. Fay (Eds.), *Development of the auditory system* (pp. 146–192). New York: Springer.
- Qi, L., Liu, H., Lutfy, J., Funnell, W. R. J., & Daniel, S. J. (2006). A nonlinear finite-element model of the newborn ear canal. *Journal of the Acoustical Society of America*, 120, 3789–3798.
- Qi, L., Funnell, W. R. J., & Daniel, S. J. (2008). A nonlinear finite-element model of the newborn middle ear. *Journal of the Acoustical Society of America*, 124, 337–347.
- Raphael, Y., Marshak, G., Barash, A., & Geiger, B. (1987). Modulation of intermediate-filament expression in developing cochlear epithelium. *Differentiation*, 35, 151–162.
- Ren, T. & Nuttall, A. L. (2001). Basilar membrane vibration in the basal turn of the sensitive gerbil cochlea. *Hearing Research*, 151, 48–60.
- Romand, R. (1987). Tonotopic evolution during development. *Hearing Research*, 28, 117–123.
- Rosowski, J. J., Carney, L. H., & Peake, W. T. (1988). The radiation impedance of the external ear of cat: Measurements and applications. *Journal of the Acoustical Society of America*, 84, 1695–1708.
- Ruah, C. B., Schachern, P. A., Zelterman, D., Paparella, M. M., & Yoon T. H. (1991). Age-related morphologic changes in the human tympanic membrane. A light and electron microscopic study. *Archives of Otolaryngology, Head & Neck Surgery*, 117, 627–634.
- Ruggero, M. A., Robles, L., & Rich, N. C. (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *Journal of Neurophysiology*, 68, 1087–1099.
- Ryan, S., & Piron, J. (1994). Functional maturation of the medial efferent olivocochlear system in human neonates. *Acta Oto-Laryngologica*, 114, 485–489.
- Sanchez-Fernandez, J. M., Rivera, J. M., & Macias, J. A. (1983). Early aspects of human cochlea development and tectorial membrane histogenesis. *Acta Oto-Laryngologica*, 95, 460–469.
- Saunders, J. C., Kaltenbach, J. A., & Relkin, E. M. (1983). The structural and functional development of the outer and middle ear. In R. Romand & M. R. Romand (Eds.), *Development of auditory and vestibular systems* (pp. 3–25). New York: Academic Press.
- Shaw, E. A. G. (1988). Diffuse field response, receiver impedance, and the acoustical reciprocity principle. *Journal of the Acoustical Society of America*, 84, 2284–2287.
- Shaw, E. A. G., & Teranishi, R. (1968). Sound pressure generated in an external-ear replica and real human ears by a nearly point source. *Journal of the Acoustical Society of America*, 44, 240–249.
- Shera, C. A. (2003). Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. *Journal of the Acoustical Society of America*, 114, 244–262.
- Shera, C. A. (2004). Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. *Ear and Hearing*, 25, 86–97.
- Shera, C. A., & Abdala, C. (2011). Otoacoustic emissions – mechanisms and applications. In K. Tremblay & R. Burkard (Eds.), *Translational perspectives in auditory neuroscience*. (In press), San Diego: Plural.
- Shera, C. A., & Guinan, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. *Journal of the Acoustical Society of America*, 105, 782–798.

- Shera, C. A., Talmadge, C. L., & Tubis, A. (2000). Interrelations among distortion-product phase-gradient delays: Their connection to scaling symmetry and its breaking. *Journal of the Acoustical Society of America*, 108, 2933–2948.
- Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2010). Otoacoustic estimation of cochlear tuning: Validation in the chinchilla. *Journal of the Association for Research in Otolaryngology*, 11, 343–365.
- Siegel, J., & Kim, D. (1982). Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity. *Hearing Research*, 6, 171–182.
- Simmons, D. D. (2002). Development of the inner ear efferent system across vertebrate species. *Journal of Neurobiology*, 53(2), 228–250.
- Smurzynski, J. (1994). Longitudinal measurements of distortion-product and click-evoked otoacoustic emissions of preterm and full-term infants. *Ear and Hearing*, 14, 258–274.
- Smurzynski, J., Jung, M. D., Lafreniere, D., Kim, D. O., Kamath, M. V., Rowe, J. C., Holman, M. C., & Leonard, G. (1993). Distortion-product and click-evoked otoacoustic emissions of preterm and full-term infants. *Ear and Hearing*, 14, 258–274.
- Song, L., McGee, J., & Walsh, E. J. (2008). Development of cochlear amplification, frequency tuning, and two-tone suppression in the mouse. *Journal of Neurophysiology*, 99, 344–355.
- Stevens, J. C., Webb, H. D., Smith, M. F., & Buffin, J. T. (1990). The effect of stimulus level on click evoked oto-acoustic emissions and brainstem responses in neonates under intensive care. *British Journal of Audiology*, 24, 293–300.
- Strickland, E. A., Burns, E. M., & Tubis, A. (1985). Incidence of spontaneous otoacoustic emissions in children and infants. *Journal of the Acoustical Society of America*, 78, 931–935.
- Talmadge, C. L., Tubis, A., Long, G. R., & Piskorski, P. (1998). Modeling otoacoustic emission and hearing threshold fine structures. *Journal of the Acoustical Society of America*, 104, 1517–1543.
- Tanaka, K., Sakai, N., & Terayama, Y. (1979). Organ of corti in the human fetus. Scanning and transmission electronmicroscope studies. *Annals of Otology*, 88, 749–758.
- Tubis, A., Talmadge, C. L., Tong, C., & Dhar, S. (2000). On the relationships between the fixed- f_1 , fixed- f_2 , and fixed-ratio phase derivatives of the $2f_1-f_2$ distortion product otoacoustic emission. *Journal of the Acoustical Society of America*, 108, 1772–1785.
- Walsh, E. J., McGee, J., McFadden, S. L., & Liberman, M. C. (1998). Long-term effects of sectioning the olivocochlear bundle in neonatal cats. *The Journal of Neuroscience*, 18, 3859–3869.
- Zweig, G. (1976). Basilar membrane motion. *Cold Spring Harbor Symposia on Quantitative Biology*, 40, 619–633.
- Zweig, G. (1991). Finding the impedance of the organ of Corti. *Journal of the Acoustical Society of America*, 89, 1229–1254.
- Zweig, G., & Shera, C. A. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *Journal of the Acoustical Society of America*, 98, 2018–2047.

Chapter 3

Morphological and Functional Development of the Auditory Nervous System

Jos J. Eggermont and Jean K. Moore

1 Introduction

There are three essential aspects to consider in the description of human auditory maturation: the structural, the functional, and the behavioral. Structural aspects can be studied by histological methods, that is, by cell and axon staining in the brains in deceased infants and children. Structure can also be studied in living persons by neuroimaging methods that visualize the density of gray and white matter [magnetic resonance imaging (MRI)] and can trace fiber tracts through the diffusion of water along or perpendicular to them [diffusion tensor imaging (DTI)]. Functional imaging methods quantify the brain's use of either oxygen or metabolites [positron emission tomography (PET), single positron emission computed tomography (SPECT)] or changes in the amount of oxygenated blood [functional MRI (fMRI), blood oxygen level-dependent (BOLD) response]. Alternatively, neural function can be assessed functionally by auditory evoked potentials (AEPs) or magnetic fields that quantify changes in polarity, amplitude, and latency, as well as localization of putative equivalent dipoles. The increasing use of modern imaging methods (MRI, DTI, fMRI) may offer views of maturation complementary to those offered by histology and electrophysiology, and these new findings are extensively covered in this chapter. Behavioral studies quantify auditory discrimination, from simple threshold measurements to speech discrimination under various conditions as a function of age, but are covered

J.J. Eggermont (✉)

Department of Physiology and Pharmacology, Department of Psychology,
University of Calgary, Calgary, AB, Canada
e-mail: eggermon@ucalgary.ca

J.K. Moore

House Ear Institute, Los Angeles, CA, USA
e-mail: drjkm@gmail.com

in other chapters of this volume. Ideally, structural, functional, and behavioral methods of assessment should give similar time lines of auditory maturation, but there may not be perfect correspondence between the results of the different approaches.

In this overview of the development of hearing, readers are reminded of the three main gradients of maturation in the auditory system. The first is the peripheral to central gradient, meaning that even if central structures mature faster, peripheral immaturity will be the limiting factor. This gradient is characterized by early maturation of the brain stem and reticular activating system (RAS) pathways, followed by a later and very extended maturation of thalamocortical and intracortical connections (Moore and Linthicum 2007). It is, however, possible that the specific lemniscal and extralemniscal auditory pathways mature at different rates than the nonspecific RAS. The parallel processing in these three pathways may offer ultimately a top-down influence on processing of auditory information (Kral and Eggermont 2007). The second gradient is the middle-to-apical and middle-to-basal cochlear maturation, also known as “the sliding place principle” (Rubel et al. 1984). Because this will be relevant only in the preterm and in the very first post-term months, one expects to see it reflected in frequency-dependent maturation of the ABR. The third gradient is that of the maturation of cells and axons in cortical layers, initially in layer I, followed by layers IV–VI and then upward to the superficial layers II–III, a result of the developmental gradient in the cortical plate (Moore and Guan 2001).

The peripheral to cortical maturational process suggests that the roughly two decades of human auditory maturation can be divided into several periods dictated by structural or functional temporal landmarks. A recent review (Moore and Linthicum 2007) divided anatomical development into a perinatal period (third trimester to 6 months postnatal), early childhood (6 months to 5 years), and late childhood (5–12 years). Although it is obvious from behavioral measures that a fetus can hear from about the 27th week, an important question is whether term birth [38–42 weeks conceptional age (CA)] is a maturational landmark for auditory function, and whether preterm birth has an effect on subsequent maturation. Although the division into several periods seems obvious (Moore 2002; Eggermont and Ponton 2003), it is perhaps more instructive to explore a separation of the maturational sequence as resulting from two major auditory processes: discrimination and perception.

The first developmental period is manifested by early auditory discrimination, and is a reflection of maturation in the brain stem and cortical layer I. This process is determined by increasing axonal conduction velocity and is largely complete at age 1 year, though fine-tuning occurs into the second year of life. One of the main puzzles in current descriptions of human auditory development is the apparent discrepancy between structural indices of relatively late maturation of the auditory cortex and the functional and behavioral results that suggest very early auditory processing capacities. Infants younger than the age of 6 months have the ability to discriminate phonemic speech contrasts in nearly all languages, a capability they later lose when raised in a one-language environment. In contrast, the histology of the brain in the first half-year of life indicates only a poor and very partial maturation of the auditory cortex. This discrepancy suggests either that infants rely largely on subcortical processing for this discrimination, or that the methods used in

quantifying the structural and physiological properties of the auditory system are incomplete, or at least insensitive.

The second major maturational period reflects the development of auditory perception, the attribution of meaning to sound, with its neural substrate in cortical maturation. This process depends on synapse formation and increasing axonal conduction velocity, and has a maturational onset between 6 months and 1 year. The age of 6 months is a behavioral turning point, with changes occurring in the infant's phoneme discrimination. This is more or less paralleled by regressive changes in the constituent makeup of layer I axons in the auditory cortex and the onset of maturation of input to the deep cortical layers. One could entertain the idea that, at about 6 months of age, the cortex starts to exert its modulating or gating influence on sub-cortical processing via efferents from the maturing layers IV–VI, resulting in the loss of discrimination of foreign language contrasts. The period between 2 and 5 years of age, the time of development of perceptual language, is characterized by a relatively stable level of cortical synaptic density that declines by 14 years of age (Huttenlocher and Dabholkar 1997). In later childhood, a continued improvement of speech in reverberation and noise, and of sound localization, is noted. At the end of the maturational timeline, one usually considers the hearing of young adolescents as completely adult-like. However, speech perception in noisy and reverberant acoustic environments does not mature until around age 15.

Maturation of auditory anatomy and behavior is reflected in progressive changes in electrophysiological responses. At approximately 2 years of age, the electrophysiological measures of auditory function in the form of the auditory brain stem response (ABR), middle latency response (MLR), the late cortical P2 component of the obligatory cortical evoked potentials (OEP), and the mismatch negativity (MMN) are fully mature. At about 6 years of age, the long-latency (~100 ms) N1 component of the OEP is typically not recordable with stimulus repetition rates above 1/3 s. Reliability improves over the next 5 years, and the N1 is detectable in all 9- to 10-year-olds at stimulus rates of approximately 1/s. Age 12 and up is characterized by major transient changes in the cortical evoked potentials that are likely related to the onset of puberty (Ponton et al. 2000a), and functional aspects of this perceptual process continue to change well into adulthood. Though behavioral measures of auditory perception are adult-like by age 15, the maturation of long-latency AEPs continues for at least another 5 years thereafter. This may suggest the need for additional behavioral studies in adolescents, and provides yet another example of the relative strengths and weaknesses of alternative methods in the evaluation and interpretation of human auditory maturation. Further, not all detectable electrophysiological or structural changes need to be behaviorally relevant.

Obviously, maturation of the central auditory system is driven by sensory input. The effects of hearing loss on auditory system maturation have been studied in children who received a cochlear implant a variable time after the onset of deafness. Maturation in the presence of partial and complete deafness is typically delayed and incomplete when the duration of the deprivation is long and occurs during early childhood, a finding supported by strong evidence from electrophysiology. The positive effects of hearing aids, and in particular, cochlear implants, are a basis for discussion of ways to ameliorate this abnormal maturational process.

2 Neural Correlates of the Discrimination System Maturation

In terms of auditory capability, it is obvious that infants in the months before and after birth respond to environmental sounds and accurately distinguish different sounds. In particular, infants younger than the age of 6 months can discriminate individual speech phonemes in both their native language and in languages to which they have not been exposed (Trehub 1976) (Table 3.1). Electrophysiological testing reveals that infants exhibit AEPs arising both in the brain stem and in the forebrain. These behavioral and physiological findings are in basic agreement with anatomical and imaging studies showing early and rapid maturation of auditory structures in the brain stem and in some elements of cortex. In this section, we review the auditory brain stem response (ABR), the middle latency response (MLR), the early-maturing obligatory evoked potentials (OEP) components, and the mismatch response (MMR), as well as the structures that are their presumed generators.

Table 3.1 Structural, electrophysiological, and behavioral correlates in maturation

Age	Structural	Electrophysiological	Behavioral
<6 months	Only layer I cortical axons are mature	ABR wave I mature	Discrimination of speech sounds, detection of change
	Acoustic radiation myelination begins	ABR waves III–V, MLR, P2 and MMR present but immature	
6 months–5 years	Brain stem axons are all mature	ABR, MLR, P2, N2, MMN, and T-complex mature	Onset and development of perceptual language
	Acoustic radiation myelination matures	P1 present but immature	
	Cortical layers IV, V and VI axonal neurofilaments mature		
	Cortical synaptic density peaks		
5–12 years	Cortical layers II and III axonal neurofilaments mature	P1 matures, N1 emerges but remains immature	Processing of masked and degraded speech improves
	Decrease in cortical synaptic density		
>12 years	Cortical axons are all mature	Asymptotic maturation of N1	Speech in reverberation and noise matures
	Temporofrontal language-related nerve tracts mature		

2.1 Brain Stem Responses and Their Generators

2.1.1 Brain Stem Structural Maturation

Histological Studies of Brain Stem Development

The brain stem auditory system is a pathway consisting of axonal tracts interrupted by synaptic relays in nuclei as depicted in Fig. 3.1. The brain stem auditory nuclei and axonal pathways are formed very early in development, being identifiable by the 8th week of embryonic life. In their earliest developmental stages, auditory axons are connected to their targets but are extremely thin and not capable of conducting action potentials. The first step in axonal maturation in the brain stem pathway is formation of the neurofilaments that are the inner framework, or cytoskeleton,

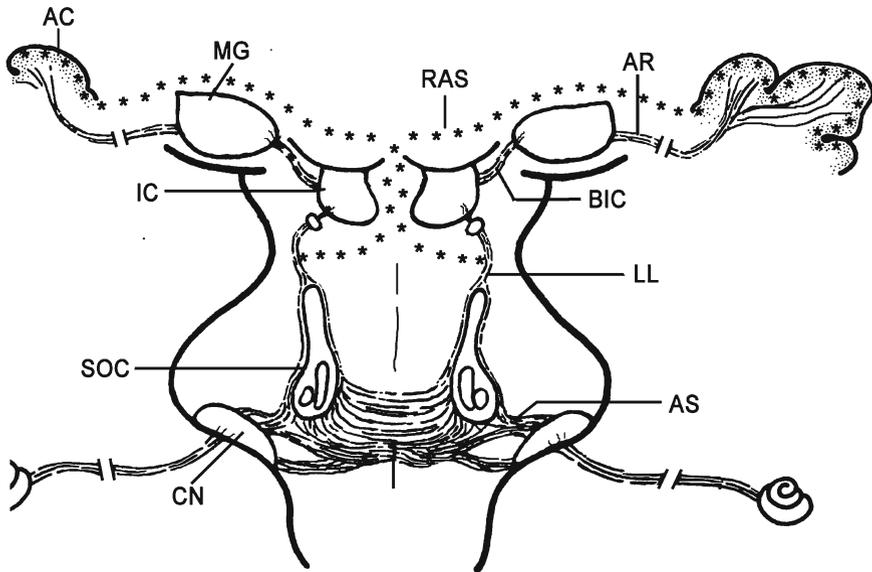


Fig. 3.1 Schematic representation of the human auditory system, with cellular structures labeled on the *left* and axonal pathways on the *right*. Nerve fibers carrying activity from the cochlea enter the cochlear nuclei (CN) and terminate on its various cell groups. Projections from the cochlear nuclei pass through the acoustic stria (AS) to innervate the nuclei of the superior olivary complex (SOC). Other axons continue through the lateral lemniscus (LL) to end in the inferior colliculus (IC). Collicular projections pass through the brachium of the inferior colliculus (BIC) to reach the auditory part of the thalamus, the medial geniculate (MG). From the thalamus, projections pass through the auditory radiation (AR) to the auditory cortex (AC) of the temporal lobe. In an alternate pathway, collaterals from lemniscal axons enter the reticular core of the brain stem, from which arises the reticular activating system (RAS) pathway to layer I of the cortex [Adapted from Fig. 1 in Moore, J.K. and Linthicum Jr., F.H. Auditory system, Chapter 34, pp. 1241–1279, in Paxinos, G. and Mai, J.K. (eds) *The Human Nervous System*, 2nd edition, 2004, with permission from Elsevier]

of the axon. As this process occurs, the total number of neurofilaments and the relative number of their light, medium, and heavy subunits controls the size of the developing axon, and accordingly, its future conduction velocity (Hoffman et al. 1984; Xu et al. 1996). In immunohistochemical studies of the human brain stem, stainable neurofilaments are first observed in brain stem auditory axons at the 16th fetal week and become more numerous by the 22nd week (Moore et al. 1997). By the 27th–29th weeks, the beginning of the third trimester, the axonal pathways have an adult-like fasciculated configuration and a network of terminal branches within their target nuclei. The next stage of axonal development is the formation of myelin sheaths. The first light myelin coatings are visible in the cochlear nerve, acoustic stria, lateral lemniscus, and brachium of the inferior colliculus by the 27th–29th week. Subsequently, visible myelin density increases steadily to at least 1 year postnatal (Moore et al. 1995). It is noteworthy that at all stages of maturation, the development of neurofilaments and myelin is equivalent at all levels of the brain stem; that is, the brain stem auditory pathway matures as a unit from the cochlear nerve to the thalamus.

Neuronal maturation within the brain stem auditory pathway is also reflected in the development of the dendritic arbors that form the basis for synaptic networks. The intracellular framework of dendrites consists of long microtubules, bound together and stabilized by cross-bridges formed by a molecule called microtubule-associated protein (MAP2) (Weisshaar et al. 1992). With antibody staining (Moore et al. 1998), MAP2 can be visualized in the cell bodies of brain stem auditory neurons at the 20th–22nd fetal weeks, but at this age there are essentially no dendritic processes. By the 24th–26th weeks, approaching the time of onset of auditory function, there are still only short dendritic branches. However, by 1 month postnatal, dendritic branches are greatly elongated, and by 6 months of age, dendrites are extensively branched and exhibit mature features such as terminal tufts. Dendritic development thus takes place mostly during the perinatal period of emerging auditory function, coincident with the process of axonal myelination.

Brain Stem Imaging Studies

Myelin maturation in the brain stem pathway has been investigated in living subjects by MRI imaging. Using signal changes in T1- and T2-weighted MRI, Sano et al. (2007) traced myelin changes in the cochlear nucleus, superior olivary nuclei, lateral lemniscus, and inferior colliculus in 192 infants and children ranging in age from 4 weeks prenatal to 224 weeks postnatal. They found that axons in the cochlear nucleus and superior olivary complex showed myelinated intensity change from 3 to 13 postnatal weeks on T2-weighted images. The lateral lemniscus showed myelinated intensity change from 3 to 8 corrected postnatal weeks on T1-weighted images, and from 1 to 13 corrected postnatal weeks on T2-weighted images. The inferior colliculus showed intensity change from 2 to 39 corrected postnatal weeks on T2-weighted images. The onset of these changes lagged 11 to 18 weeks behind histological changes, indicating less sensitivity of MRI to early stages of myelin

formation, but allowing detection of continued maturation in the inferior colliculus after the lower brain stem appeared mature.

Thus histological techniques appear to allow visualization of early changes in small numbers of elements, but imaging seems better able to track the rate and continued course of maturational changes.

2.1.2 Brain Stem Electrophysiological Activity

The scalp-recorded response to activation of axons in the auditory nerve and brain stem is called the ABR. A typical ABR consists of a sequence of up to seven vertex-positive waves (recorded from the top of the head), separated by negative valleys. The peaks are typically labeled with roman numerals, while the valleys are generally not labeled. It is known that wave I is the compound action potential of the initial part of the auditory nerve. Our interpretation of the ABR morphology generated by more central structures has been considerably aided by several important findings in the 1980s, including correlation of surface activity with depth recordings during surgery (Hashimoto et al. 1981; Møller et al. 1988) and modeling experiments (Stegeman et al. 1987). The work of Stegeman and co-workers showed that recordable far field potentials at the scalp originate from spatially stationary voltage peaks in the brain stem. These stationary peaks only arise when action potential activity (1) travels from one medium into another with a different conductance; (2) experiences a change in geometry such as branching, or (3) traverses a bend in the nerve tract. Thus, it is likely that wave II is generated by the auditory nerve at the point where it leaves the petrous bone to enter the intracranial space. Peaks with numbers from III to V likely are generated sequentially in the auditory brain stem, but waves IV and V may also reflect parallel activation of ipsilateral and contralateral pathways. Wave III is likely generated by nerve branching within the cochlear nucleus. Wave IV is most likely generated by axons passing directly from the dorsal cochlear nucleus (DCN) and ventral cochlear nucleus (VCN) to the contralateral lateral lemniscus, as the track bends in passing around the superior olivary complex. Wave V probably arises from fibers synapsing in the medial superior olive (MSO) and running into the lateral lemnisci on both sides of the brain toward the inferior colliculus.

The ABR can be reliably recorded in premature infants from the 28th–29th weeks CA (Starr et al. 1977; Ponton et al. 1992; Hafner et al. 1991). The studies of Pasman et al. (1991) found that at 30–35 weeks CA, the ABR vertex–ipsilateral mastoid wave I, the negative wave following wave II, wave V, and the vertex contralateral mastoid recorded waves II and V were the most consistently present with detection rates of 87–100% (Fig. 3.2). At ages 35 weeks CA and older, waves I, III, and V at both sides were clearly present (Pasman et al. 1991; Ponton et al. 1992).

Clearly, ABR maturation will be affected by cochlear maturation. Although the cochlea has a generally adult appearance by the end of the second trimester, investigations utilizing distortion product otoacoustic emissions (DPOAE) and cochlear traveling wave delay measurements indicate that full cochlear maturity is not achieved until a few weeks before term birth (see Abdala and Keefe, Chap. 2),

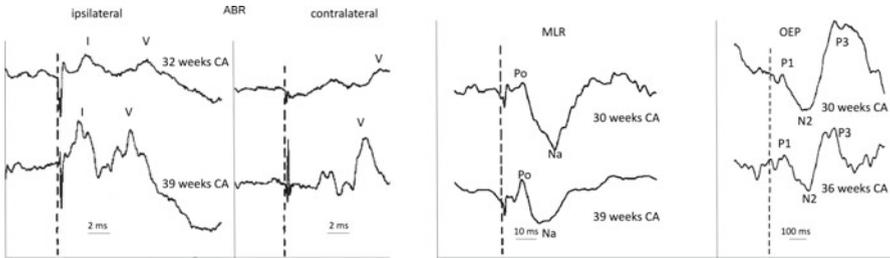


Fig. 3.2 Prenatal changes in ABR (ipsi- and contralateral), MLR and late cortical potentials at 30–32 weeks (*top row* in each panel) and around term. Some of the peaks have been identified. Note the dramatic changes in the ABR waveforms where contralateral activity lags behind the ipsilateral one, and the much more modest ones for the OEPs [Redrawn from Pasman et al. 1991]

with an additional 3–6 months required for maturation of hair cell-auditory nerve fiber synapses mediating frequencies above 6,000 Hz (Eggermont et al. 1991, 1996). Thus, just as in the output of the cochlea is represented by wave I, all subsequent ABR generators show a cochlear place-dependent maturation, characterized by its most sensitive frequency [characteristic frequency (CF)] (Ponton et al. 1992). Using click stimulation and high-pass noise masking, combined with exponential curve fits for the latency-age data, it has been shown that the wave I–V latency difference, termed the brain stem conduction time, exhibits the most rapid maturation for input from the middle section of the cochlea. This section is formed by the octave frequency bands centered on 1.4, 2.8, and 5.7 kHz, and its I–V interval matures with a time constant of 29 weeks, that is, every 29 weeks the I–V interval difference with the adult value is reduced by a factor 2.72 (see Appendix for more details). For all practical purposes, the maturational process underlying the I–V interval reaches adult values after two or three time constants have elapsed (decrease in I–V interval difference with the adult value by a factor 7.4–20), in this case 58–87 weeks after the first recordings at 35 weeks CA, which is about 1–1.5 years post term. For the apical part of the cochlea, represented by the octave band centered at 0.7 kHz, the time constant of I–V interval maturation is about 48 weeks, and for the basal part of the cochlea, that is, the frequency band centered around 11.3 kHz, the time constant is 40 weeks. This converts into maturity at approximately 2–2.5 years and 1.5–2 years post term, respectively. Using unmasked clicks, the time constant of maturation of the I–V interval is about the same as that for the middle frequency range and equals about 30 weeks, thus maturing at 1.5 years of age. This suggests that when using unmasked-click stimuli, one probes the most mature part of the system.

The ABR potential maturational time course, reflecting activity in brain stem structures up to the level of the superior olivary complex, is closely followed by the MLR. The MLR components that are most clearly detectable in infants are the P0–Na complex (Fig. 3.2). Because recordings from the surface of the human brain stem match the P0 peak to a postsynaptic potential at the inferior colliculus (Hashimoto et al. 1981), it seems likely that the P0–Na waves reflect transmission in the brachial pathway from the inferior colliculus to the thalamus (Fig. 3.1). The P0–Na waves are barely detectable at the 25th–27th fetal weeks, but are fairly

well defined by the 33 rd fetal week, and more pronounced by the time of term birth (Rotteveel et al. 1985, 1987; Pasman et al. 1991). The Na peak latency decreases steadily from about 28 ms at the 30th fetal week to approximately 20 ms at term (Rotteveel et al. 1985, 1987; Pasman et al. 1991). By the third postnatal month, the Na peak achieves a latency of approximately 18 ms, a value that remains unchanged throughout childhood, teen years, and adulthood (Kraus et al. 1985).

2.1.3 Structure–Function Relationships

Neurofilament formation increases the intra-axonal diameter, while myelin formation, occurring on the outer surface of the axons, provides electric insulation of the axon. These two structural processes jointly determine axonal conduction velocity. The described synchrony in the development of myelin in cochlear nerve and brain stem implies that rapidly conducted axonal potentials appear at about the same time in pathways from the cochlea to the inferior colliculus and in the collicular projection to the medial geniculate body. Potentially then, the onset of rapidly conducted action potentials should occur at much the same time along the entire pathway from the cochlea to the thalamus.

If age-related latency shifts are, in fact, due to progressive myelination of the auditory tract, then there should be a general correspondence between the process of myelination and the time course of ABR latency changes during the perinatal period. In studies of premature infants, the observed rate of change in interwave latencies is quite large between the 28th and 40th weeks (Starr et al. 1977; Ponton et al. 1992). In postnatal infants, the values of the interwave intervals decline more slowly and approach adult values by 1–2 years of age (Lieberman et al. 1973; Mochizuki et al. 1983; Jiang et al. 1991; Ponton et al. 1992). It thus appears that the changes in myelin density and axonal velocity run generally parallel, with rapid change from 30 to 40 weeks of gestation and slower change through the first and second postnatal years (Table 3.1). Similar early and rapid maturation of latencies is observed in the P0–Na complex generated in the upper brain stem.

In regard to conduction velocity, increasing myelin thickness should mean faster conduction, causing decreasing ABR peak latencies. It has been shown that the ABR intervals reflecting only axonal conduction time (I–II and III–IV) have an adult-like conduction time by the time of term birth (Moore et al. 1996). However, at the time of term birth, adult-like conduction time may not mean adult-like conduction velocity, because the brain stem is growing and thus the auditory pathway is still lengthening. For instance, the distance between cochlear nucleus and the upper contralateral lateral lemniscus increases from 12 mm at 20 weeks CA to approximately 36 mm at 90 weeks CA (~1 year of age), and to about 41 mm in the adult. In the same period the conduction velocity in this tract increases from 6 m/s to about 35 m/s (adult), which compensates nearly perfectly for the increase in path length. It thus appears that increasing myelin thickness, by compensating for increasing pathway length, keeps conduction time in pathways without synapses steady in the first year of life. Synaptic changes, therefore, appear to be responsible for late-stage maturation in brain stem conduction time.

2.2 *Early-Maturing Event-Related Potentials (Fields) and Their Generators*

2.2.1 **Early Structural Maturation of the Auditory Cortex**

Histological Studies of Cortex Development

The basic structure of auditory cortex matures rapidly in the perinatal period. Modern investigations (Moore and Guan 2001) are in agreement with the earlier extensive studies of Conel (cited in Moore and Guan) in concluding that cortical neurons enlarge and develop their dendritic arbors during these months (Fig. 3.3c). This cell growth and differentiation determines that cortical cytoarchitecture (i.e., cortical depth and lamination) is adult-like by age 1–2. In contrast to cortical layers II–VI, which are made up of tightly packed cell bodies, layer I consists mostly of cell processes, that is, axons and dendrites. A prominent dendritic component of layer I is the terminal tufts of apical dendrites of cells in deeper layers, predominantly those in layers II, III, and V (Fig. 3.3a, c).

In the months before and after term birth, axonal maturation in cortex is confined to layer I. One axonal system, a layer of thick axons at the base of the layer, arises from a group of intrinsic neurons, the Cajal-Retzius (C-R) cells (Fig. 3.3b) (Marin-Padilla and Marin-Padilla 1982). These cells can be identified in human layer I by their staining for acetylcholine esterase (Meyer and González-Hernández 1993) and calretinin (Meyer and Goffinet 1998; Spreafico et al. 1999; Zecević et al. 1999). Though small-molecule amino acid transmitters are too labile to be identified in human postmortem material, the excitatory transmitter glutamate can be demonstrated in C-R cells of animals (del Rio et al. 1995). Neurofilament staining of human cortex indicates that C-R axons progressively develop their capacity for conduction, from a small number of axons at the 22nd fetal week to a prominent population at 4.5 months postnatal (Fig. 3.4; Moore and Guan 2001). The excitatory influence of this system may, however, be modulated by a population of non-C-R cells that contain the inhibitory transmitter γ -aminobutyric acid (GABA) (Imamoto et al. 1994).

In addition to C-R cell intrinsic axons, dendritic tufts of pyramidal cells in layer I are contacted by afferent axons coming from lower levels of the nervous system. Around the time of birth, these very thin axons enter the cortex from below and form branches within layer I that run tangentially, for distances up to several millimeters (Fig. 3.3a). These axons are present as early as the 7th gestational week (Marin-Padilla and Marin-Padilla 1982), and are presumed to come from the earliest developing part of the central nervous system, the brain stem reticular formation (Fig. 3.1). This input is commonly referred to as the RAS. As with the intrinsic C-R axons, neurofilament staining shows an increasing number of mature axons from the time of term birth to at least 4.5 months postnatal (Fig. 3.4; Moore and Guan 2001). The organization of this system, with its long tangential axons contacting the dendrites of large numbers of cortical neurons, would promote broad activation of the neurons in auditory cortex. However, these thin and very lightly myelinated

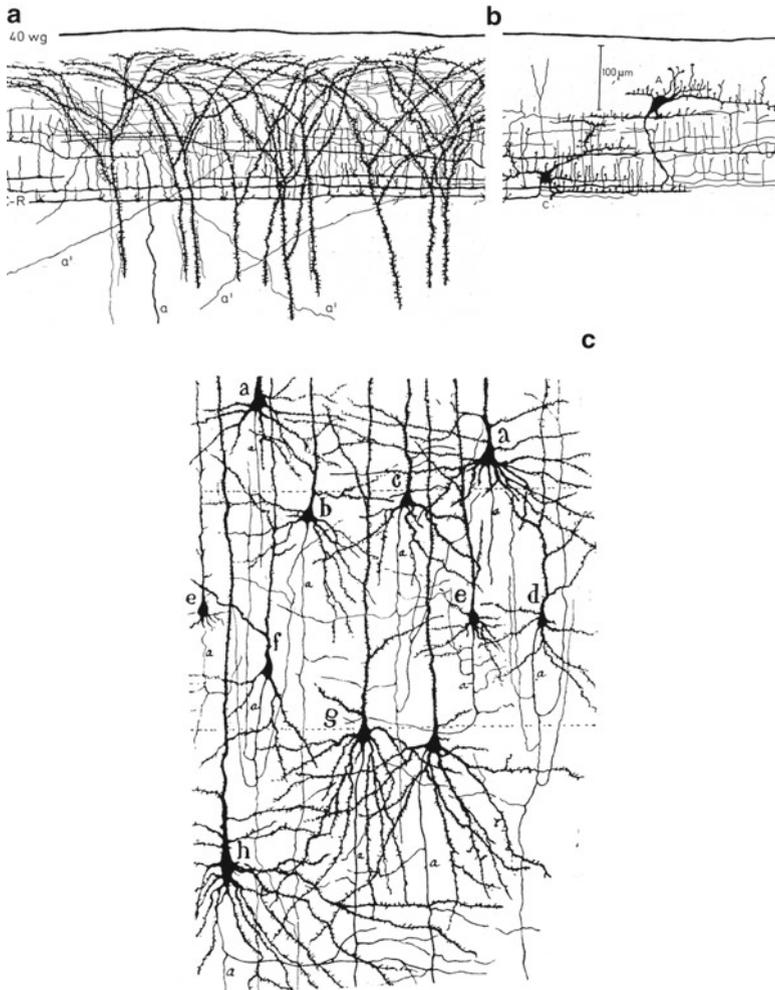


Fig. 3.3 Neuronal development of the auditory cortex in the human newborn. **(a and b)** components of layer I: **(a)** apical dendritic tufts of neurons in deeper layers, crossed by thin brain stem axons (a') and thick axons of Cajal-Retzius cells; **(b)** intrinsic Cajal-Retzius cells and their axons; **(c)** pyramidal neurons in layer V with well developed basal and apical dendrites. [**a and b** adapted from Figs. 4a–c and 5a in Marin-Padilla and Marin-Padilla (1982), with kind permission from Springer Science+Business Media. **c** adapted from Fig. 7 in Ramon y Cajal (1900) (in Spanish)]

axons must have a very slow conduction velocity, and therefore could contribute only to long-latency evoked potentials.

In the second half of the first year of life, changes occur in layer I that are transitional to the cortical organization seen in older children and adults. Between 4.5 months and 1 year of age, there is a complete disappearance of the layer of C-R axons (Fig. 3.4; Moore and Guan 2001), consistent with the fact that C-R cells

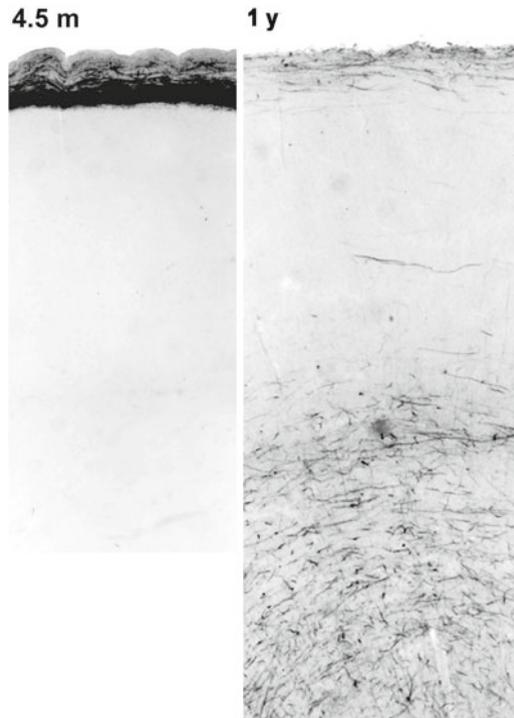


Fig. 3.4 Development of axonal neurofilaments in auditory cortex in human infants. At 4.5 months, both the superficial (reticular) and deep (Cajal-Retzius cell) tiers of layer I axons are well developed, but no maturing axons are seen in the deeper cortical layers. By 1 year of age, there has been a reduction in number of the thin extrinsic axons and a complete loss of the thick C-R cell axons in layer I. In addition, axons with mature neurofilaments are beginning to appear in the deeper cortical layers [Adapted from Figs. 5 and 7 in Moore and Guan (2001), with kind permission from Springer Science + Business Media]

undergo postnatal apoptosis (i.e., controlled degeneration) both in animals (Imamoto et al. 1994; del Rio et al. 1995; Hestrin and Armstrong 1996) and in humans (Spreafico et al. 1999). The GABAergic intrinsic cells, however, remain present into adulthood (Winer and Larue 1989). During this same period, the number of RAS axons in layer 1 is markedly reduced (Fig. 3.4). By 1 year of age, along with regression of the C-R and RAS systems, a number of relatively thick spiral axons are observed running upward to reach and enter layer I (Moore and Guan 2001). These ascending axons in human cortex appear homologous to axons identified in monkeys as projections from the medial (magnocellular) division of the medial geniculate body (MGm) (Burton and Jones 1975; Hashikawa et al. 1995), and thus would provide the first source of input from the thalamus to layer I. These thalamic axons presumably account for the numerous asymmetric-membrane, round-vesicle excitatory synapses observed on dendritic spines within layer I of adult animals (Beaulieu and Colonnier 1985). Because this system of input travels through the main auditory pathway to the thalamus, instead of through the thin-axon RAS

pathway, it has the potential to generate shorter latency layer I–evoked potentials. Because it arises from a nontonotopically organized part of the medial geniculate complex, it is part of the extralemniscal pathway.

At the same time, during the first postnatal year, there is a gradual onset of development of thalamocortical projections to deeper layers of cortex (Fig. 3.4). The acoustic radiation (Fig. 3.1) is a large bundle of myelinated axons running from the posterior thalamus to the auditory cortex in the upper temporal lobe (Bürgel et al. 2006). By 1 month after birth, some neurofilament maturation is evident in axons of the acoustic radiation (Moore and Guan 2001). This filament maturation is followed by incipient myelination in the radiation from 3 to 4 months postnatal, as detected by histological staining (Yakolev and Lecours 1967; Kinney et al. 1988) and structural MRI (Pujol et al. 2006). This gradually developing “bottom-up” mode of stimulation of auditory cortex neurons likely interacts with their “top-down” activation by layer I axons, with resulting modification of the early-maturing cortical potentials.

Functional Imaging of Auditory Cortex

Studies employing fMRI show maturation of the cortical response to auditory stimuli in infants with increasing age. In neonates, activation of temporal cortex is inconsistent, in that of 20 neonates exposed to tonal stimuli, six showed no response, nine showed a decrease in blood flow, and five showed increased blood flow (Anderson et al. 2001). The polarity of the blood oxygen level–dependent (BOLD) signal depends on the amounts of oxygenated and deoxygenated hemoglobin present in tissue volumes. In adults and older children, the increase in fMRI signal is thought to arise from the disproportionate increases in cerebral blood flow and oxygenation relative to oxygen consumption by the brain. In the younger infants, the negative signal change may be due to the inability of the cerebral vasculature to meet increased oxygen demand by increasing local blood flow, leading to a decrease in the ratio of oxyhemoglobin to deoxyhemoglobin. However, by 3 months of age, all infants demonstrate a positive BOLD response to speech stimuli (Dehaene-Lambertz et al. 2002). In both studies, the area activated was the superior temporal plane bilaterally, greater on the left, and with the highest level of response, whether positive or negative, in the primary auditory area within the transverse temporal gyrus. The Dehaene-Lambertz study also demonstrated activity in higher-level auditory areas such as the superior temporal and angular gyri. Thus, overall, the pattern of cortical regional activation evoked in infants via layer I stimulation is remarkably similar to the pattern seen in older children and adults.

2.2.2 Electrophysiological Measures of Cortical Activity

In perinatal infants, brain stem potentials are not the only sign of auditory activity in the central nervous system, given that cortical potentials can be recorded as well. Cortical potentials present in infants include both the longest-latency OEPs and the MMR. Both of these types of potentials undergo changes across the first years of postnatal life.

The Mismatch Response

Presentation of a novel (oddball or deviant) stimulus produces a larger event-related response (ERP) compared to that for a frequent stimulus just preceding it. The difference between the oddball and frequent ERP is called the mismatch negativity (MMN). The presence of an MMN has been seen as an indicator of preattentive detection of stimulus change, be it acoustical, phonetic, or contextual (Näätänen 2001). One has to realize that the MMN is never recorded as such from the scalp; it is a construct designed by the investigator and is obtained by subtracting two nonsimultaneously recorded ERPs. Thus, in interpreting the MMN one should always inspect the individually recorded OEPs, or their magnetic field equivalent, the obligatory magnetic fields (OMFs) that are at the basis of this construct.

The MMN is considered an ERP related to the suppression of activity of the frequent stimulus as a result of its (quasi-) periodic presentation. In infants, a particular maturational OEP sequence is observed: The predominantly negative cortical OEP waveform as observed in adults has been found to be dominantly positive in neonates and infants (Fig. 3.2). This positivity is first present over the frontocentral region around term, and later, by 1–2 months of age, is also present over the temporal region. In infants, the early difference response is of positive polarity, and to avoid confusion, it is often called the mismatch response (MMR). The suppression of the frequent-OEP is clearly present in 3-month-old infants (Dehaene-Lambertz and Gliga 2004) and can form the basis for finding an MMR in infants and even in fetuses as young as 28 weeks gestational age (Draganova et al. 2005, 2007). Infants of 8 months of age still showed slow positive MMR to/da/-/ta/ phoneme contrasts (Pang et al. 1998), whereas using gap duration as the contrast, a transition to an adult-like MMN at 6 months of age was observed (Trainor et al. 2003), suggesting that the MMR to different stimulus contrasts matures at different times (He et al. 2007). In term infants, the initial positivity recorded from midline electrodes and the negativities recorded from ipsi- and contralateral temporal electrodes did not correlate in their peak and offset latencies, suggesting independent generators for each of these components (Novak et al. 1989).

He et al. (2007) investigated the emergence of discriminative responses to pitch by recording 2-, 3-, and 4-month-old infants' OEP responses to frequent and infrequent pitch changes in piano tones (Fig. 3.5). In all age groups, the infants' responses to deviant tones were significantly different from the responses to the frequent tones, suggesting that the two tones were processed as different. However, two types of MMR were observed simultaneously in the constructed difference waves. By low-pass filtering of the MMR, an increase in the left lateralized positive slow MMR emerged that was prominent in 2-month-olds, present in 3-month-olds, but insignificant in 4-month-olds. By high-pass filtering a faster, more adult-like MMR, lateralized to the right hemisphere, emerged at 2 months of age and became earlier and stronger with increasing age. The coexistence but spatial dissociation of the two types of MMR suggested different underlying neural mechanisms.

The infant cortex may differentially process incoming information, as infants are prone to alternate sleep-wake cycles. During sleep, the positive MMR dominates,

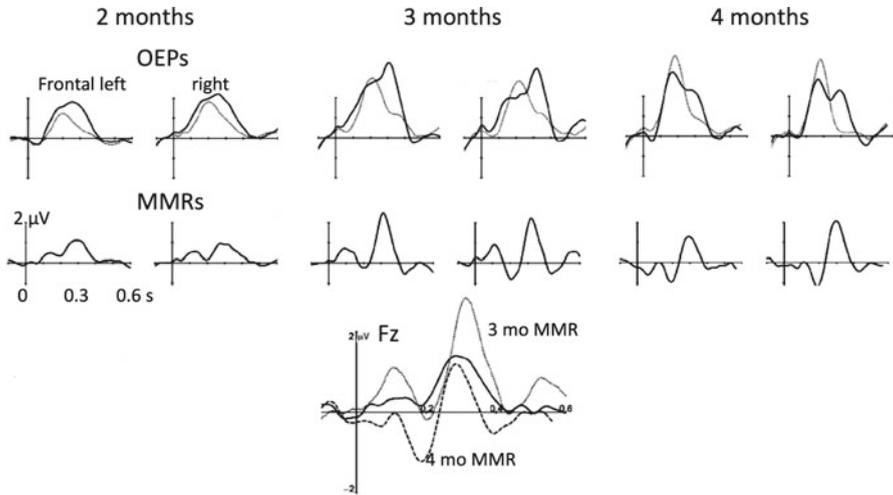


Fig. 3.5 OEPs to frequent (*dotted lines*) and deviant (*solid lines*) stimuli are shown in the *top layer panels* for 2-, 3-, and 4-month-olds. The *middle layers* show the corresponding MMRs. In the *bottom central panel*, the MMRs for the Fz recording site are shown. FL and FR stand for *left and right frontal region*, which bracket the midline location Fz [Redrawn from He et al. 2007]

whereas the negative polarity was more obvious in waking infants (Friederici et al. 2002). Thus there may be two overlapping mismatch responses (Kushnerenko et al. 2002): a negative MMR that is later than the MMN of an older child or adult, and a positive MMR. The negative wave may represent the detection of the deviant, and the positive wave may reflect attention to this stimulus, not unlike the P3a component of the mature ERP. Depending on the relative timing and sizes of these waves, a negative or positive MMR might result (Picton and Taylor 2007). Again, one has to keep in mind that the MMR does not exist in the brain, but rather is artificially constructed. One can only argue that activity in one part of the brain is going more negative while another part is going more positive, and that the distribution of activity across cortical layers and regions determines the summed changes observed in scalp recordings.

2.2.3 Obligatory Evoked Potentials and Fields

The aforementioned findings all suggest that auditory cortex metabolism and electrical or magnetic evoked activity are related to the behavioral change detection capabilities of infants. Because MMRs are derived from OEPs (or OMFs), this also implies that the OEPs are detectable at an early age. Early studies (Barnet et al. 1975; Ohlrich et al. 1978; Rotteveel et al. 1987; Pasman et al. 1991) showed that in preterm babies at 24 weeks the OEP is dominated by a large negative wave with a peak latency of 200 ms, followed by a positive peak at 600 ms that by 30 weeks is reduced to a latency approaching 300 ms and increases in amplitude (Fig. 3.2). By the time of

term birth, this positive peak has a latency of about 250 ms and dominates the response because the earlier negative wave has nearly completely disappeared.

This dominance of positive components continues until about age 5 months, when a new negative peak with latency of approximately 400 ms begins to increase in amplitude and reduce in latency. Even by age 5 years, this sequence of positive (100 ms)–negative (200 ms)–positive (~350 ms) peaks is still very much the standard morphology for OEPs recorded at the vertex (Pang and Taylor 2000; Ponton et al. 2000a). Barnet et al. (1975) showed that during the first 3 years of life the latencies of the three dominant waves—P2, N2, and P3—decreased linearly with the logarithm of the age. Eggermont (1988) showed that an exponential decrease with age (time constant of ~40 weeks) comprehensively described the changes in P2 and N2 latency and was interpretable in biological terms, whereas the often-used log-age representations were not (Ponton and Eggermont 2006; see also the Appendix). This time constant is only slightly longer than that of the maturation of the wave I–V latency difference in the ABR, suggesting that the maturation of this early cortical activity is either limited by that of the auditory brain stem or “co-tuned” to it. Because the RAS pathway originates as an offshoot of the brain stem pathway, getting its auditory information from collaterals of lateral lemniscus axons, brain stem pathway activity would definitely be a synchronizing factor.

2.2.4 Structure–Function Relationship

There is one set of data that seems to contradict the idea that the onset of auditory function occurs, as measured by ABR and behaviorally, most commonly in the 27th–29th weeks of development. This contradictory material is represented by cortical potentials evoked with click stimuli as early as the 25th fetal week (Weitzman and Graziani 1968). In fact, when cortical potentials and ABRs were obtained from the same population of premature infants, the cortical potentials usually appeared earlier (Starr et al. 1977). These observations are puzzling in light of the observations that myelination does not occur in brain stem pathways until after the 26th–28th weeks and, more pertinent, does not occur in thalamocortical projections until after the 40th week of development (Langworthy 1933; Yakolev and Lecours 1967; Kinney et al. 1988). The most likely explanation for this phenomenon is cortical excitation through RAS pathways arising within the core of the brain stem. The lower brain stem is always the first step in processing, giving rise to both the RAS pathway and the thalamocortical pathways. In addition, the biphasic action potentials that form the ABR waves do not synchronize well in infancy because of slow-conduction velocities in the unmyelinated or poorly myelinated brain stem tracts, whereas the much slower postsynaptic potentials that form the cortical OEPs do this adequately, despite being also activated by slowly conducting fibers. Thus the cortical maturational process, as witnessed through electrophysiological means, probably reflects a change in balance between the RAS system and the various thalamocortical pathways.

The next problem that arises is the very early maturation of the MMR as an indicator of change detection. What neural structures are used to accomplish this, and what

underlies the change from the slow positive MMR to the faster negative-positive MMR at around 4–8 months of age depending on the acoustical change contrast? In 2-month-old sleeping infants, the OEPs to frequent stimuli are large and positive and the response to the deviant is larger and also positive, yielding a positive MMR. In awake 2-month-old infants, the OEPs to frequent and deviant stimuli are both smaller but still positive, with the latter showing a negative rebound at very long latencies (600–800 ms; Friederici et al. 2002). Positive MMRs, resulting from positive obligatory OEPs, suggest a depolarization site (sink) in layer IV, the entrance of thalamocortical activation, which is likely the earliest specific or lemniscal auditory input to the cortex. Myelination of the acoustic radiation begins only at 3 months, so if these unmyelinated fibers are functional, the results could be long-latency OEPs and MMRs. Alternatively, there could be an active hyperpolarization site (source) in layer I, where the axons are clearly mature (see Eggermont 2006 for a discussion of sinks and sources). At 4 months of age, the OEPs are still broad positive peaks for both frequent and deviant stimuli but the differences are that the OEP to the deviant is smaller for short latencies and larger for long latencies (Fig. 3.5). Thus the MMR has become a biphasic negative–positive wave, at least for frontal and central electrodes (He et al. 2007). In the late preteen years, the MMR becomes determined by the difference between the now dominant negative N1 components of the OEP for the frequent and the deviant condition (Gomot et al. 2000). The “real” MMN, then, is clearly the result of changes in a superficial late depolarization (sink) in layer I and/or upper layer II (Javitt et al. 1996).

In a study of identified neurons responding to horizontal layer I inputs (Cauller and Connors 1994), all cells that responded were pyramidal neurons with cell bodies in layers II, III, and V and apical dendrites extending into layer I. This implies that a subset of cortical neurons may be specialized to engage the horizontal layer I inputs selectively. In particular, the apical dendrites of cells in layers IV and VI do not reach layer I, and these cells are typically the target of the thalamocortical inputs. This suggests a division of labor for lemniscal versus extralemniscal and RAS inputs to the pyramidal cells in cortex. In contrast to the excitatory postsynaptic potential–inhibitory postsynaptic potential (EPSP–IPSP) sequence characteristically evoked by deep layer electrical stimulation, horizontal layer I inputs in adults evoked long-lasting EPSPs (~50 ms). In many layer V neurons, the initial EPSP evoked by horizontal layer I stimulation was followed by a variable late depolarization that was blocked by the *N*-methyl-*D*-aspartate (NMDA) receptor antagonist 2-amino-5-phosphonovaleric acid (APV) (Cauller and Connors 1994).

As discussed in the Sect. 2.1.1, on histological studies, cortical neurons are strongly driven during the perinatal months by the glutamergic, and thus excitatory, C-R cell system in layer I. The rapid involution of the C-R cell population in the middle of the first year of life ushers in a period in which GABAergic, non-C-R cells may exert a stronger, unopposed influence. By age 1 year, thalamic (MGm) axons are beginning to form the more mature excitatory innervation of layer I. During this same period, from 3 to 4 months of age, the acoustic radiation has begun to myelinate, a process that continues until 4–5 years of age. Thus, from about 3 months of age, some synchronized activation of layer IV synapses should be expected.

Based on all of the above, a set of mechanisms (I refer again to Eggermont (2006) for details about sinks and sources) that could accommodate the various MMR and obligatory OEPs observed in the preterm and postnatal period is the following:

1. During the perinatal period, there are functional glutaminergic C-R synapses in layer I on dendrites of pyramidal cell in layers II, III, and V. Because these synapses release glutamate, they affect the slow NMDA receptors. Activation of these synapses will induce a layer I sink (depolarizing synapse; extracellular negative potential) that will be visible as a scalp negative deflection. This can explain the large, broad, and very long latency negative OEPs recorded in preterm infants.
2. After 4.5 months of age, glutamate C-R cells disappear while GABA cells remain and now only cause transient hyperpolarizations in layer I synapses. This results in layer I sources that are visible as scalp-positive deflections, which could explain the emerging positive obligatory OEPs and consequently the resulting positive MMRs. It should be noted that a positive MMR generated by this mechanism would require increasing hyperpolarization for the deviant stimulus. Alternatively, positive MMRs can be derived from increased positive obligatory OEPs resulting from a depolarization site (sink) in layer IV, the entrance of (unmyelinated) thalamocortical activation, which is likely the earliest specific or lemniscal auditory input to the cortex.
3. With increasing maturity, other largely excitatory inputs may arrive in layer I from the MGm, which will again produce layer I sinks that would give rise to negative polarity MMNs. This, however, cannot explain the persistent positive polarity OEPs.
4. Slow-conducting, thalamocortical inputs that start to myelinate at 3–4 months of age could produce poorly synchronized depolarization in layer IV with the required scalp-positive OEP polarity (see under point 2). However, with the MMR resulting from a repetition-related decrease in the OEP to the frequent stimulus, this would also produce a positive MMR, which is observed rarely after the first 6 months of age. Though this is an incomplete mechanism, an interaction with the mechanism mentioned under point 3 can account for both negative MMRs and positive OEPs. This would be accomplished by partial cancellations at the scalp of the activity generated by these two mechanisms.
5. The dominance of a negative MMR after 6–8 months of age would require that, at this early age as well as in adults, it is generated in superficial layers I and/or II and results from excitatory modulations of activity in pyramidal cells with cell bodies in layers II, III, and V.

3 Neural Correlates of the Perceptual System Maturation

The transition from infancy to preschool years is a time of dramatic and dynamic neural and cognitive development. Brain volume shows a 25% increase between birth and about 4 years of age (Courchesne et al. 2000), and this increase is almost

entirely attributable to expansion of the forebrain cortex. In the maturing auditory system, cortical depth more than doubles between birth and age 1 (Moore and Guan 2001), while cortical synaptic density doubles from birth to age 6 (Huttenlocher and Dabholkar 1997). Because conscious perception is based on cortical function, cognitive growth during this time is also rapid (Table 3.1). Children progress from only a few words in their expressive vocabulary at 1 year to full sentences by 3 years of age. In the period between 6 months and 5 years, changes in speech perception occur, resulting in a bias toward the native language, such that discrimination of nonnative language sounds deteriorates (Werker and Tees 1984; Kuhl et al. 1992; Panneton and Newman, Chap. 7). This change coincides with the emergence of a differential cortical response to native and nonnative language contrasts in the MMN (Näätänen et al. 1997). Perceptual ability continues to mature gradually during school-age years. Between 5 and 12 years, perception of degraded speech and speech in noise gradually improves (Elliott 1979; Eisenberg et al. 2000). Children's consonant identification abilities reach adult-like levels of performance at about age 15 in reverberation-only and noise-only listening conditions, but identification in the reverberation-plus-noise listening condition does not mature until the late teenage years (Johnson 2000).

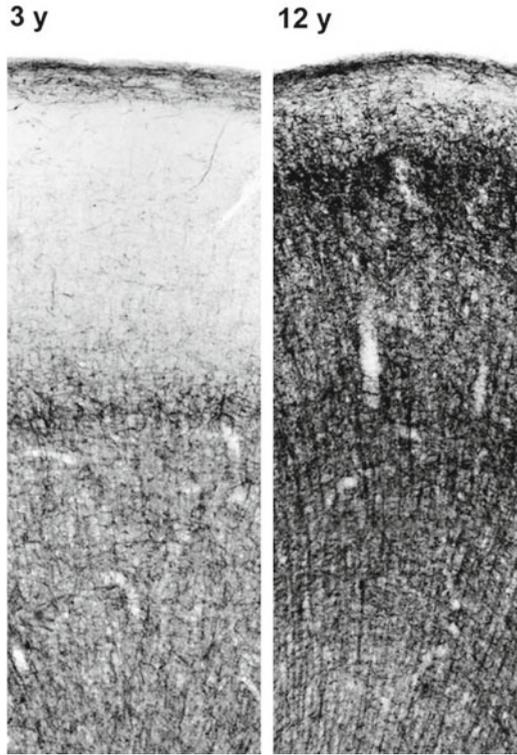
3.1 Late-Maturing Obligatory Evoked Potentials (Fields) and Their Generators

3.1.1 Continuing Structural Maturation of the Auditory Cortex

As noted in the previous section, the basic cytoarchitecture and laminar organization of auditory cortex is adult-like by the end of the first or second year of life (Moore and Guan 2001). Yet the pattern of evoked cortical potentials changes radically across infancy, childhood, and teen years. A possible explanation lies in the fact that the various systems of axons driving the cortex have a multiyear course of maturation. As previously discussed, the earliest developing system of cortical activation is layer I, driving the infant cortex, first by reticular and intrinsic axons, and later by thalamic input. During the subsequent childhood years, auditory cortex maturation consists of development of thalamocortical and intracortical axon systems, both of which are important for normal cortical function.

Thalamic axons to the cortex have a typical pattern of termination, consisting of dense innervation of layer IV, with sparser horizontal collaterals to layer V and short vertical collaterals extending into the deepest part of layer III (Pandya and Rosene 1993; Hashikawa et al. 1995). The previous section referred to development of filaments and myelin in the initial segment of auditory thalamic axons in the acoustic radiation during the first year of life. Evidence of maturation also appears in the distal part of the axons as they course through the deeper layers of cortex. Neurofilament immunostaining (Moore and Guan 2001) demonstrates small numbers of filament-stained axons traversing layers IV, V, and VI at ages 1 and 2. By age 3,

Fig. 3.6 Development of axonal neurofilaments in auditory cortex in children. By age 3 years, axons with mature neurofilaments show the typical pattern of thalamocortical afferents in the *deeper layers*, but no axonal maturation has occurred in the *upper layers*. By age 12, axons in the superficial layers have mature neurofilaments, giving the cortex an essentially adult appearance. [Adapted from Figs. 7 and 8 in Moore and Guan (2001), with kind permission from Springer Science+Business Media]



the typical pattern of large endings in layer IV and short collaterals into the deepest part of layer III is evident (Fig. 3.6). By age 5, mature thalamocortical axons fill the deeper cortical layers (Moore and Guan 2001). The next step in maturation of these thalamocortical axons, namely myelination, has been observed to occur progressively from the 3rd to 4th postnatal months until about 4–5 years of age (Yakolev and Lecours 1967). Thus in the perceptual maturation phase, the same area of temporal cortex is being activated as in the perinatal period, but the method of activation is different. In contrast to the diffuse horizontal layer I system, the thalamocortical input forms a vertical, columnar, point-to-point system of connections.

The human auditory cortex, and its thalamic input, is not homogeneous. Rather, the auditory cortex consists of areas that are distinguished by their cytoarchitecture (Brodmann 1908; Galaburda and Sanides 1980). The primary or core auditory area (Brodmann area 41) is located on the upper surface of the temporal lobe, in the transverse temporal gyrus (Fig. 3.7). Surrounding the core area is a belt of secondary auditory cortex (area 42). Extending anteriorly, posteriorly, and laterally over the surface of the superior temporal gyrus is a parabelt region of higher level auditory processing (area 22). Studies in primates have demonstrated that the tonotopically organized lemniscal pathway from the ventral nucleus of the medial geniculate (MGv) projects only to the core and belt areas of cortex (Burton and Jones 1975;

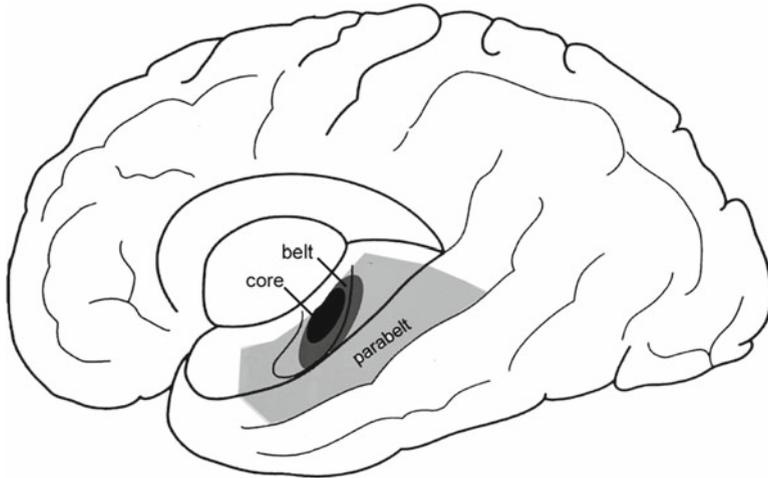


Fig. 3.7 *Left hemisphere, with the lateral fissure opened to reveal the superior surface of the temporal lobe and the transverse temporal gyrus. Shading illustrates the core and belt areas of auditory cortex on the superior temporal plane and the parabelt area extending laterally and down over the superior temporal gyrus*

Aitkin et al. 1988; Rauschecker et al. 1997; de la Mothe et al. 2006). A partially overlapping set of extralemniscal projections from the medial geniculate complex innervates the belt and parabelt more diffusely. This broad projection includes non-tonotopic projections from the magnocellular nucleus of the medial geniculate (MGm), sparse input from other ancillary nuclei (dorsal medial geniculate, suprageniculate/posterior, limitans), and extensive input from the medial part of the pulvinar (Trojanowski and Jacobson 1975; Aitkin et al. 1988; Hackett et al. 1998a; Luethke et al. 1989; Rauschecker et al. 1997; de la Mothe et al. 2006). Though the axons arise from different sources, the course of maturation of the thalamocortical axons innervating the core, belt, and parabelt is identical across the time period from 3 to 6 months to 4–5 years (Moore and Guan 2001).

A second stage of cortical maturation occurs in later childhood. By age 5–6 years, mature thalamocortical axons fill the deeper cortical layers, but axons in layers II and III still show little sign of stainable neurofilaments (Moore and Guan 2001). However, by age 11 or 12, the density of mature axons in the upper layers has become equivalent to that of an adult (Fig. 3.6). Studies in primates have shown that layers II and III are the source and target of many of the association axons running between the core and belt (Aitkin et al. 1988; Rauschecker et al. 1997), as well as between the belt and parabelt (Hackett et al. 1998b). In addition, most of callosal neurons connecting the auditory cortex with the corresponding area in the contralateral hemisphere arise and terminate in layers II and III. (Aitkin et al. 1988; Luethke et al. 1989; Pandya and Rosene 1993; Hackett et al. 1999). Thus maturation of layer II and III axons significantly broadens the scope of intracortical processing, both within and between hemispheres.

3.1.2 Imaging Studies of Auditory Cortex

Structural Imaging

Diffusion tensor imaging (DTI), a recent MRI technique that assesses water diffusion in biological tissues at a microstructural level, has become a powerful technique to explore the structural basis of brain development (Hüppi and Dubois 2006). Water diffuses preferably along myelinated axon tracts. Parameters that allow an assessment thereof are the mean diffusivity and the fractional anisotropy (FA); the latter is based on diffusion differences in the two directions. These parameters are calculated in each voxel of the image. D_{av} is the mean of the parallel and perpendicular diffusion values. During white matter development, D_{av} reflects premyelination changes in axonal width and axon neurofilaments, and in myelination. The increase in white matter FA values during development also takes place in two steps. The first increase again takes place before the histological appearance of myelin. The second, more sustained increase in anisotropy is associated with the histological appearance of myelin and its maturation. This two-stage increase in white matter anisotropy occurs at different rates (and thus different time constants) for different brain areas (Hüppi and Dubois 2006). An overview is presented in Table 3.2.

In the thalamus of subjects between 5 and 30 years old, single exponential maturation time courses were found both for D_{av} and for FA, with time constants in the range of 5.6 years and 8.9 years, respectively. The 90% maturation milestone was reached between 13 and 20 years (Lebel et al. 2008).

In the temporal and frontal regions of cortex, some myelinated white matter is seen relatively early, but quantitatively relevant changes are observed only after the fifth month. Half of the children in a recent study (Pujol et al. 2006) showed a myelinated white matter content of 10% by the age of 18 months in language-related areas. A total of 90% of children met this myelination criterion for language-related regions only after 35 months. White matter relative volume changes could be fitted by a single exponential approach to the adult value with a time constant of about 30 months. This suggests that temporal cortex myelination is adult-like at about 100 months of age, that is, at approximately 7–8 years, comparable to the maturation of white matter in the thalamus and around the time that the N1 component of the AEP appears. This estimate is smaller by a factor of about 2 than that obtained from a cohort of 88 volunteers between 3 and 30 years old (Pfefferbaum et al. 1994; Paus et al. 2001). This is to be expected given the nonoverlapping and very different age groups; the larger the age span the longer the time constant is in case of prolonged maturation. Maturation of FA with a time constant equal to 29 months was found in a group ranging in age from neonates to 10-year-old children for the superior longitudinal fasciculus (SLF) the major temporofrontal tract, which plays an important role in language. This maturation time course parallels that of myelination that also occurs in the first 2 years of life (Zhang et al. 2007). In contrast, Lebel et al. (2008) described maturation in the superior longitudinal fasciculus (SLF) for subjects between 5 and 30 years old as much slower and very similar both for D_{av} and FA with time constants of 7.5 years and 9.6 years, respectively, again similar to

Table 3.2 Maturation along the auditory system

Structure	Test	τ Fast	τ Slow	τ	Maturity	Authors
Auditory nerve	ABR wave I			4 weeks	3–6 months	Eggermont et al. 1991
Brain stem	ABR-click	4 weeks	59 weeks		2 years	Eggermont 1988
Brain stem	ABR-mid freq			29 weeks	2 years	Ponton et al. 1992
Thalamus	DTI	8 weeks	9 years		20 years	Mukherjee et al. 2001
Thalamus	DTI			7 years	15–20 years	Lebel et al. 2008
Acoustic radiation	Histology				4–5 years	Kinney et al. 1988
Auditory cortex	MRI			5.5 years	12–16 years	Su et al. 2008
Temporal cortex	DTI			2.5 years		Pujol et al. 2006
SLF	DTI			2.5 years		Zhang et al. 2007
SLF	DTI			8.9 years	13–20 years	Lebel et al. 2008
T complex, P2, Pa, Pb	AEP			40 weeks	1.5–2 years	Eggermont 1988
N2	AEP			2 years	5 years	Ponton et al. 2000a
P1, N1, TP200	AEP			4–9 years	15–20 years	Ponton et al. 2002

the findings for the thalamus. The 90% milestone was reached between 13 and 20 years. These contrasting findings, albeit obtained in nonoverlapping age groups, still suggest two very distinct stages of maturation. In a T2-weighted MRI study comprising 241 neonates and young children up to 8.5 years, Su et al. (2008) observed myelination changes in auditory cortex to progress with a time constant of 5.5 years. However, in children 5 years of age and older no significant differences from adults were detectable.

Functional Imaging

Using single photon emission computed tomography (SPECT) to measure regional cerebral blood flow (rCBF) and cerebral glucose metabolism (rCGM) in regions associated with higher-order language skills, Devous et al. (2006) found that the primary auditory area had become a stable neuronal population by age 7 years. In contrast, ongoing maturation (reflected by declining rCBF and attributed to dendritic pruning) continued into adolescence in higher order cognitive areas, such as the angular gyrus. Similar results were obtained when developmental differences in brain activation of 9- to 15-year-old children were examined by fMRI (Cone et al. 2008). During an auditory rhyme decision task to spoken words, a developmental increase in activation in the left superior temporal gyrus (BA 22) was seen across all lexical conditions, suggesting that automatic semantic processing increases with age regardless of task demands.

3.1.3 Electrophysiological Measures of Cortical Function

Across early childhood, there is increasing prominence of short-latency over longer-latency positive cortical OEP components, which may be related to the newly maturing system of thalamocortical connections. The middle latency components typically mature earlier than the long-latency OEPs, consisting of the P1–N1–P2–N2 complex with mature latencies of about 50, 100, 150, and 200 ms. The adult waveform P1–N1–P2–N2 complex (Fig. 3.8) is achieved between 14 and 16 years of age (Pasman et al. 1999; Ponton et al. 2000a).

P1 can be distinguished at 30 weeks CA as a broad wave with a latency of 80–100 ms (Pasman et al. 1991). Over time, its peak latency gradually shortens to reach an adult level of 50 ms. The late appearance of N1 is partly due to its greater sensitivity to interstimulus interval (ISI) in children than in adults. Therefore most studies using ISIs shorter than 1 s do not observe N1 reliably before age 9 (Ponton et al. 2000a), but with a longer ISI, N1 become visible from age 6 on (Paetau et al. 1995; Gomes et al. 2001; Gilley et al. 2005). Although the latency changes for P1 and N1 peaks are similar, the maturational changes in magnitude are opposite; P1 magnitude decreases while N1 increases with increasing age. Because N1 emerges in the OEP at about 9–10 years of age, when the neural generators producing the P1 peak are essentially adult-like, and given the partial temporal overlap and common

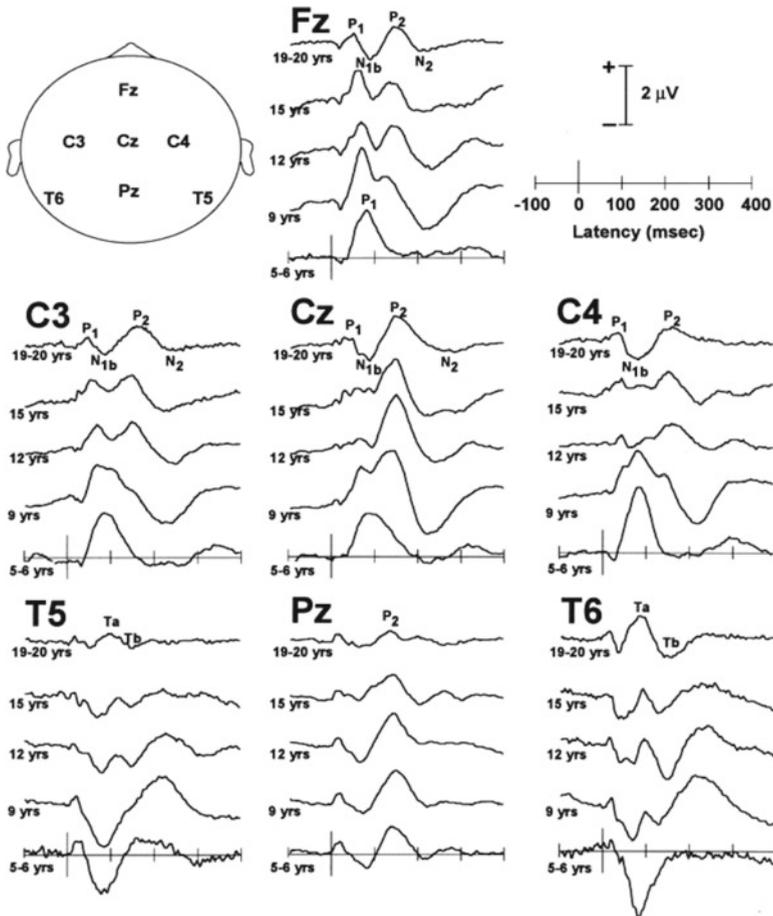


Fig. 3.8 Age-dependent morphology of the OEPs for different recording sites. Note the late (~9 years) appearance of N1 in the Fz and Cz recordings [Reprinted from Ponton et al. (2000a), with permission from Elsevier]

tangential dipole orientation of these two components, it is possible that the magnitude and latency changes of the maturing N1 peak are superimposed on those of P1 (Ponton and Eggermont 2001).

During early maturation it is difficult to identify clearly a component as N2 when the N1 is not yet visible. However, their maturational trajectories are different, as shown by the N1/N2 amplitude ratio increase with age. The N2 is mature much earlier (time constant 2 years) than the N1 (time constant >4 years), as shown in Ponton et al. (2002). Further, the scalp distributions of both the N1 and N2 were found to change with maturation, which might indicate different component structure (and thus function) of children's and adults' N1 and N2 (Ceponiené et al. 2002).

Results of dipole source modeling (Ponton et al. 2002) demonstrated that the three orthogonal dipole components of the sources in each hemisphere isolate three distinct sets of OEP components. The MLR peaks Pa and Pb are best represented by the sagittal (anteroposterior) dipole sources; the “classic” P1–N1–P2–N2 sequence is isolated to the tangential (to the scalp) sources that are perpendicular to the superior surface of the temporal lobe; and the T-complex peaks Ta and Tb, together with the TP200, are represented in the radial (perpendicular to the scalp) dipole sources. The grouping of OEP components isolated in each orthogonal dipole remained the same across a 5- to 20-year age span. This suggests that the orientations of the OEP generators are essentially adult-like by 5 years of age. But note that Albrecht et al. (2000) attribute the emergence of the N1/P2 complex in adolescence (in children older than 8 years) to a change of the tangential dipole source potential. However, one has to be careful to use the term “N1/P2 complex,” as P2 matures early (reaching adult values by as early as 2–3 years of age), while the N1 follows a much longer developmental time course, extending into adolescence (Ponton et al. 2000a).

Based on a global measure of similarity between individuals’ OEP waveforms, Bishop et al. (2007) distinguished three developmental periods: 5–12 years, 13–16 years, and adulthood. There was a fairly sharp change in the OEP waveform at Fz around 12 years of age, when N1–P2 was fully mature. Ponton et al. (2000a, 2002) distinguished three maturation groups of potentials: one group reaching maturity at age 5–6 and consisting of the MLR components Pa and Pb, the OEP component P2, and the T-complex, recorded from temporal electrodes. The temporally recorded Ta/Tb complex, which is distinct from N1, is already mature at age 5–6 (Pang and Taylor 2000; Gomes et al. 2001; Tonnuquist-Uhlen et al. 2003). Because of its radial-oriented dipole it is likely generated in BA22. A second group that was relatively fast to mature (time constant 2 years) was represented by N2 only. A third group was characterized by a slower pattern of maturation (time constants of 4–9 years) and included the OEP components P1, N1, and TP200, the long-latency component following the T-complex. The observed latency differences combined with the differences in maturation rate indicate that P2 is not identical to TP200. The results also demonstrated the independence of the T-complex components, represented in the radial dipoles, from the P1, N1, and P2 components, contained in the tangentially oriented dipole sources.

3.1.4 Structure–Function Relationship

The work of Moore and Guan (2001) on the structural maturation of the human auditory cortex, and the sink-source description of synaptic activation (Mitzdorf 1985), have led to a straightforward interpretation of the cortical layer of origin of the scalp-recorded AEPs. Moore and Guan (2001) studied cortical structural maturation using immunostaining of axonal neurofilaments, which are the structural scaffolding of axons. In an immature state, axon diameter is small, and because conduction velocity is related to diameter, these smaller-diameter axons have slower conduction velocities. As neurofilaments mature, axons grow to their full diameter, allowing action potentials to be conducted at a velocity of approximately 1 m/s.

Fast conduction velocities mean that activity carried by individual neurons from various locations (thalamus, other cortical areas) arrives synchronously and results in sharp, short-latency and large-amplitude AEPs. This synchronization is also a prerequisite for adequately processing the temporal aspects of communication sounds (Michalewski et al. 2005). Below 4.5 months of age, only axons in superficial layer I of the auditory cortex are immunostained and can thus be considered mature; the absence of axonal neurofilament staining elsewhere in the cortex indicates immaturity of the axons in all deeper layers. Maturation of axons in the deeper layers occurs progressively from 1 to 5 years of age, and of axons in the superficial layers between 5 and 12 years.

From the polarity of the OEP components, one has to conclude that scalp-positive components such as the P1 and P2 are generated in the lemniscal input layers (lower III–IV) of auditory cortex. P1 is likely mature by age 12 (Ponton and Eggermont 2001), comparable to the time of maturation of the thalamus as shown in DTI (Table 3.2). This was estimated from cochlear implant cases who failed to develop an N1. As most scalp-positive evoked potential components originate from excitatory synaptic activity in layers lower III and IV, most of these (e.g., the middle latency responses) indeed appear fully mature with respect to latency and amplitude by that age. The pedunculo-pontine tegmental (PPT) nucleus, a cholinergic subdivision of the reticular formation that receives auditory input, may be significant for generation of the human P1 (Harrison et al. 1990). An overview of time constants and age of maturity is given in Table 3.2.

In contrast, the scalp-negative components must originate from excitatory inputs in more superficial layers (II, upper III). This is based on the location of the activated excitatory synapses that generate these components in animals (Eggermont and Ponton 2002; Eggermont 2006). Several studies have suggested that the development of the N1 component may relate to the formation of functioning synaptic connections within the upper layers of the auditory cortex (Ponton et al. 2000a, 2002), with the synaptic activation (sinks) in older infants and children concentrated in the deeper layers of the cortex (causing a surface positive wave) and not extending fully into the upper layer parts of the dendritic trees (which would generate a surface negative wave). Because of this prolonged maturation of superficial cortical layers, it is not surprising that, for stimulus repetition rates of ≥ 1 Hz, N1 cannot be recorded below the age of 8–9 years. The maturation of N1 extends well into adolescence and appears to be associated with activity originating from the newly mature upper layer II synaptic activity.

However, this explanation of late detectability of N1 does not at first sight fit with the early emerging responses in the same latency range recorded from the lateral scalp, the so-called T-complex (Bruneau et al. 1997; Tonnquist-Uhlen et al. 2003), which is likely generated in auditory association regions of the cortex (Ponton et al. 2002). The T-complex consists of two surface negative waves, Na and Tb, even in infants with latencies in adulthood similar to N1. Picton and Taylor (2007) speculated that the infant auditory response generated in the supratemporal plane contains a large positive wave with peak latency between 100 and 200 ms that obscures an also present smaller negative N1 component. However, the maturation of the T-complex is very different from that of N1, so it is unlikely that the underlying

sources would be similar. As mentioned in Tonnquist-Uhlen et al. (2003), the Tb component of the T-complex does not show further maturational changes after 5 years of age, and latency and amplitude are statistically similar to those of P2, a component that has been ascribed to RAS activation of cortex. This was similar for electrodes T4 and T6 (both contralateral to the stimulus ear). OEPs at T6 are a faithful inversion of those at C4. Whereas the overall morphology on T4 is different from a time-inverted C4, the probability that Tb is the inverse of P2 remains.

A very interesting issue is the relationship between the protracted maturation of the N1 and the development of higher cognitive skills during childhood. The N1 emerges after children have normally acquired basic skills of verbal language, whereas language acquisition is heavily contingent on auditory sensory processing (Table 3.1). This is consistent with the view that the N1 does not index the perception of the sound features but reflects facilitative or integrative processes, sound detection, and orienting (for reviews, see Näätänen and Picton 1987; Näätänen 2001).

4 Sound Deprivation and Changing Acoustic Environments

The structural basis for the effects of deprivation on auditory function has been reviewed recently in Moore and Linthicum (2007). It is clear that neurons are generated and axonal connections established very early in development, but some aspects of neuronal maturation are delayed until the time of onset of auditory function. In both brain stem and cortex, the earliest period of auditory function is the time of formation of dendritic processes and axon terminal branching, a process that establishes the normal pattern of synaptic connections. Studies in experimental animals have demonstrated that sound deprivation reduces the number of neuronal processes and distorts their normal geometry. In addition, myelination, the final stage of axonal maturation, begins coincident with the time of onset of function. In addition, animal studies show that a lower level of electrical activity in axons retards normal myelin development. Thus, overall, sound deprivation could negatively impact the maturation of both synaptic organization and axonal conduction velocity. These findings provide at least a partial explanation for the observed abnormalities in evoked potentials in deafness, as well as for the degree of recovery occurring after prosthetic restoration of sound, that is, “time-in-sound” effects.

4.1 *Effects of Deprivation on Maturation of the Discrimination System*

4.1.1 Effects of Deprivation on Brain Stem Responses

Gordon et al. (2006) studied auditory brain stem responses to cochlear implant stimulation in 75 prelingually deafened children and 11 adults. Electrically evoked

auditory brain stem response (EABR) latencies showed significant decreases in early latency waves and interpeak latencies that occurred within the first 1–2 months of implant use. For waves eN1 (the electrically evoked negative wave following the normal wave I, which is not detectable because of the stimulus artifact); eII; eIII; and interwaves eN1–eII, eN1–eIII, and eII–eIII, the associated time constant of maturation was only 2 weeks. Slower maturation was found for eV and eIII–eV, which measure activity in the more rostral brain stem. The time constant obtained for wave eV and eIII–eV maturation was 22 weeks, implying maturity at about 1–1.5 years of age. This is slightly shorter than the 29 weeks reported by Ponton et al. (1992) for normal hearing infants. It appeared that EABR latencies in children using cochlear implants were not dependent upon the duration of deafness/ age of the children at the time of implant. Comparisons of EABR and ABRs using the time-in-sound (chronological age minus duration of deafness) variable (Ponton et al. 1996a, b) suggest that much of the change in the normal ABR III–V is activity dependent.

This was further emphasized in the study of Thai-Van et al. (2007) comprising a group of children with early-onset of deafness (mostly congenital) who had undergone cochlear implantation at between 1 year 2 months and 12 years 5 months of age (mean=3 years and 4 months). A second group of children who had become profoundly deaf after 1 year of age was also studied. In this second subgroup, age at implantation ranged from 2 years to 17 years 4 months (mean=7 years and 4 months). No influence of age at implantation on the rate of wave V latency change was found. The main factor for EABR changes again was the time-in-sound. Indeed, significant maturation was observed over the first 2 years of implant use only in the group with early-onset deafness. In this group, maturation of wave V progressed as in the ABR model of Eggermont and Salamy (1988) for normal hearing children: a sum of two decaying exponential functions, one showing an early rapid decrease in latency (time constant 3.9 weeks) and the other a slower decrease (time constant 68 weeks). In contrast, relatively little change in wave V was evident in children with late-onset deafness, as it had already reached its adult latency.

The cochlear implant–evoked middle latency response (EMLR) was detected in only 35% of 50 children (age 5.3 ± 2.9 years) at initial stimulation, but became 100% detectable after at least 1 year of use. Most children (46/50) were prelingually deaf. The older the children's chronological age, the more likely it was that the EMLR was initially detected (Gordon et al. 2005). Latencies after 6 months of implant use were prolonged in the younger age group and decreased with implant use. EMLR changes with chronic cochlear implant use suggested an activity-dependent plasticity of the central auditory system. Gordon et al. (2005) suggested that eMLR evoked in only 35% of children at initial stimulation reflects abnormally poor neural synchrony in the extralemniscal pathways (at least for post-thalamic generators). The idea that these responses are generated by the extralemniscal system agrees with the predominance of the extralemniscal contribution to the MLR in young normal hearing children. Based on a lack of N1 cortical wave responses in a small group of children with cochlear implants, Ponton and Eggermont (2001) suggested that the lemniscal pathway functions abnormally after an undefined period of auditory

deprivation. Because the belt and parabelt receive their input from the core, this could mean that extralemniscal contributions to the EMLR remain dominant in children even after long periods of cochlear implant use.

4.1.2 Effects of Deprivation on Early-Developing Cortical Responses

Activation of the extralemniscal pathway appears to be generally unaffected by a period of deafness preceding cochlear implant use. The OEP components that reflect this extralemniscal pathway activation include the T-complex and the MMR. At this early age, the MMR is likely generated by the RAS-layer I pathway. As indicated by Ponton et al. (2000b), the MMR is present even in the absence of the extralemniscal-pathway generated N1.

The morphology of the long-latency OEPs was substantially altered in implanted children after at least 3 years of complete deafness by the absence of a normal N1 peak. However, the MMN was robustly present in this group of implanted children who typically became postlingually deaf and who had good spoken language perception through their device (Ponton et al. 2000b). The differential effects that early-onset profound deafness and cochlear implant use have on the maturation of the MMN and the N1 strongly suggest that these potentials originate from different central pathways. Whereas the N1, if present, is almost exclusively unilateral in implanted children, the MMN develops earlier, with a more symmetrical representation over both hemispheres. Thus, the MMN appears to have a cortical representation much more similar to the P1 peak than to the N1 peak. It is likely that the age of maturation for the MMN is the same as that for the P2. The maturation for P2 has been equated with that for the ABR, becoming adult-like at about the age of 2 years (Eggermont 1988). This early maturation could indicate a function for the MMN similar to that for the reticular activating system, which signals the presence or absence of sensory input.

4.2 Effects of Deprivation on Maturation of the Perceptual System

4.2.1 Effects of Deprivation on Later-Developing Cortical Responses

Ponton et al. (1996a, b) compared cortical evoked potentials recorded in implanted and normal-hearing children and found that age-dependent latency changes for the P1 component, fitted to a decaying exponential curve, showed the same pattern. For implanted children, however, maturational delays for P1 latency approximated the period of auditory deprivation before implantation. This correspondence suggested that the time-in-sound determined the stage in the maturational process. This indicates that the cortical auditory system does not mature without stimulation. Once stimulation is restored, however, the normal rate of maturation for this cortical

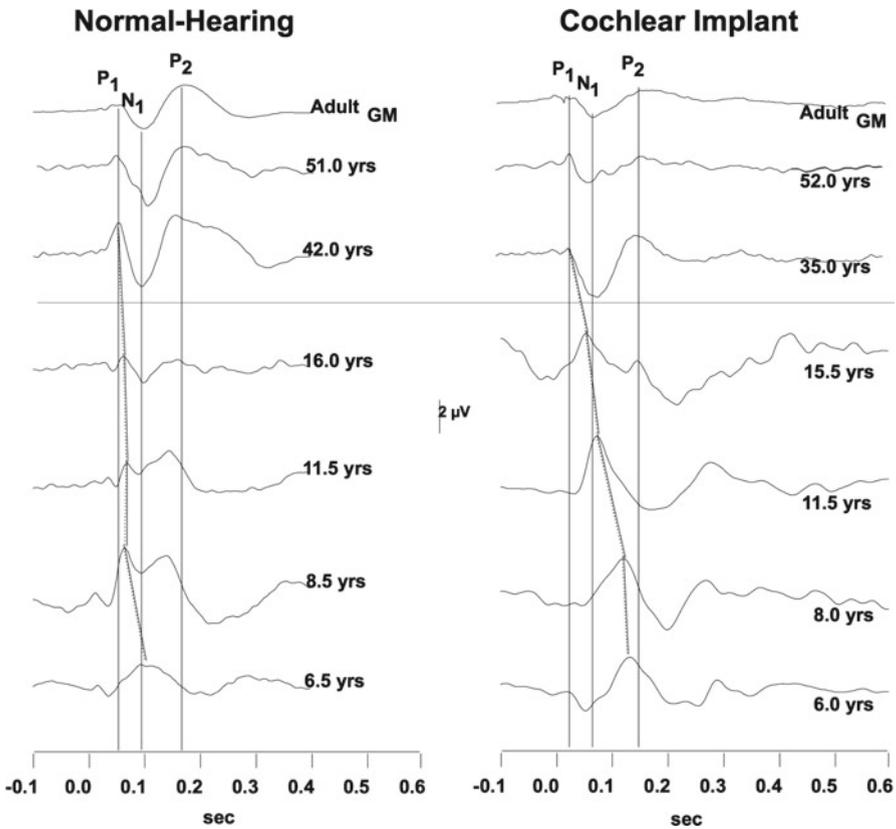


Fig. 3.9 OEP waveforms in normal hearing (*left*) and persons with cochlear implants (*right*). The waveforms in late implanted adults, which previous normal hearing are very similar to those in normal hearing adults. Note the absence of N1 in the late-implanted children. Lines connect the P1 peaks [Partially based on Ponton et al. 1996b]

activity resumes, even after an extended period of sensory deprivation. Nonetheless, the auditory system retains its plasticity during the period of deafness because the reintroduction of stimulation by the cochlear implant, perhaps as little as 50% of the time, resumes the normal maturational sequence.

However, the absence of an N1 component after implantation and the absence of any sign that it appears with long cochlear implant use (Ponton and Eggermont 2001) suggests that deprivation periods of more than 3 years, but potentially shorter, in children younger than the age of 6 are detrimental for normal maturation of auditory cortex (Fig. 3.9). In the cochlear implant users examined in this study, P1 remained much larger and broader than in their normal hearing counterparts. The P2 is present as well as a very late and broad positivity, which diminishes in amplitude with increasing age. In addition, the MLR peaks Pa, Nb, and Pb were unaffected (i.e., had normal latencies and amplitudes) by late onset (>2–3 years) auditory deprivation.

It should be recalled that the scalp-recorded OEPs represent the sum of neural activity generated in at least three central nervous system pathways: the lemniscal pathway to the primary cortical areas, the extralemniscal pathway to secondary areas, and the reticular activating system (RAS) pathway through layer I to all areas. Activation of the RAS pathway is generally unaffected by a period of profound deafness and cochlear implant use because P1 (and likely Pb) and P2 were present in the OEPs of the implanted children. These two peaks are presumed to represent activation of the RAS pathway, which consists of the ascending RAS brain stem pathway and its thalamic and cortical projections. It has been suggested that both P1 and P2 reflect the activity of a sensory gating process for auditory cortex located at the level of the reticular nucleus of the thalamus (Skinner and Yingling 1976; Yingling and Skinner 1976). The GABAergic input from the reticular nucleus blocks activity in the thalamus, but the rebound from this blocking (low-threshold calcium spikes) could result in strong activation of cortex in layer IV, the likely layer of origin for P1 and P2. Based on the presence of these peaks in the AEPs of the implanted children, the general function of this alerting or gating process appears to remain intact after a period auditory deprivation.

The absence of a normal N1 potential clearly points to abnormal function in the extralemniscal pathway, or its transcortical projections, as N1 is generated in nonprimary cortical areas, notably the planum temporale (Lütkenhöner and Steinsträter 1998). Based on the presence of Pa and the absence of N1, activation of the lemniscal pathway is unaffected by late-occurring auditory deprivation up to and including input to the deep cortical layers, which is the termination zone for thalamic cortical fibers. However, activation of superficial cortical layers is clearly affected in implanted children, and this affects transmission of activity to nonprimary areas. In contrast, the N1 is robustly present bilaterally in children affected by congenital unilateral deafness.

The major new interpretation of the abnormal maturation of AEP waveforms in implanted children proposed by Ponton and Eggermont (2001) rests on the effects that a persistent immaturity of the layer II axons, which synapse on the dendrites of the pyramidal cells, has on the generation of N1 and on the morphology of the OEPs. A simple broadening of the N1 peak by a factor of about 2–3 models most of the OEP morphology features seen in implanted children. In addition, the late positivity that is present in all children younger than the age of 8 appears to persist in implanted children. The model assumes that the maturation of all OEP components, other than N1, are unaffected by profound deafness, at least in the case of relatively late onset (>2–3 years of age). It is proposed that implanted children who experience a sufficiently long period of deafness before the age of 6–8 years never develop a fully functional set of superficial layer axons. This is supported by findings that in 104 congenitally deaf children who had been fit with cochlear implants at ages ranging from 1.3 years to 17.5 years, those with the shortest period of auditory deprivation, approximately 3.5 years or less, evidenced age-appropriate latency responses within 6 months after the onset of electrical stimulation (Sharma et al. 2002, 2007). Results of Gilley et al. (2008) based on multielectrode scalp recordings and standardized low-resolution brain electromagnetic tomography (sLORETA) imaging of

the current dipoles suggests that early implantation results in activation of the same cortical areas for P1 as in normal hearing people, whereas implantation after the age of 7 showed mainly activation in parietotemporal cortex and none in the classical auditory cortical area.

5 Summary of Anatomical–Physiological Correlations

An overview of the endpoints of the maturation process and the corresponding time constants (based on data shown in Table 3.2) are shown in Fig. 3.10 and form a maturational time line for the following summary.

5.1 Anatomical and Physiological Correlates of Discrimination

Discrimination, as a behavioral process, is the ability to recognize the differences between auditory stimuli, and involves two subsystems:

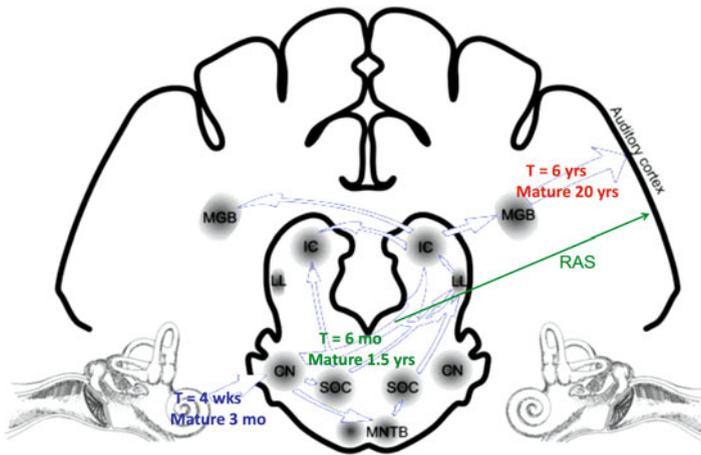


Fig. 3.10 Anatomical and functional maturation periods of the human auditory system (see Table 3.2 for individual measurements). The fastest maturing system is the cochlea and auditory nerve with a time constant of 4 weeks and reaches maturation at approximately 3 months. The brain stem up to and including the fibers into the *MGB* mature with a time constant of approximately 6 months and reach maturity at approximately 1.5 years of age. This also includes the maturation of the *RAS* pathway that innervates cortical layer I. The thalamus, the auditory radiation, and the cortex do not appear to be mature until approximately 20 years of age [Partially based on Kral and Eggermont 2007]

- Analysis of the physical parameters of the stimulus is performed by the brain stem.
- Attention to and awareness of the stimulus is mediated by the RAS (and later thalamic) input into the layer I system.

Discriminative ability and both subsystems are operational in the months before and after term birth.

5.1.1 Analysis System: Brain Stem and ABR/Po–Na

- Brain stem nuclei perform a detailed analysis of stimulus physical parameters, following the spectral analysis performed by the cochlea.
- Histological studies show rapid development of the brain stem pathway, including cells and axons, in the months before and after term birth.
- Brain stem development occurs synchronously from cochlear nerve to thalamus.
- Biological studies show that neurofilament development is autonomous, but myelination is driven by axonal activity.
- Imaging studies confirm the rapid maturation of myelin in the early postnatal months.
- ABR biphasic waves reflect activity in myelinated pathways of the lower brain stem.
- ABR development, with onset early in third trimester and rapid postnatal maturation, mirrors the anatomical development of the brain stem, with shortening latency reflecting the effect of increasing myelination on conduction velocity.
- The developmental pattern of Po–Na mirrors that of the myelination of the brachium of the inferior colliculus, in being present at birth and before, with significant maturation by age 3 months.
- Once established, there should be no change in the functioning of the brain stem pathway over the lifespan.
- Early-onset deafness affects the EABR and EMLR (Po–Na), but later-onset deafness does not, presumably because the brain stem mechanisms are already mature.
- In children with cochlear implants, EABR and EMLR maturation is activity dependent (time-in-sound), presumably because activity promotes myelination.

5.1.2 Attention System: Cortex Layer I and P2/MMR

- Attention/awareness is a function of the cortex.
- At birth and in the first months of life, the only specific (i.e., ignoring monoamine networks) pathway to cortex is the RAS system to layer I.
- RAS input consists of very thin, slowly conducting axons, which implies generation of long-latency potentials.

- Within layer I, activity is magnified by the intrinsic C-R cell axon system.
- At birth and in the first months of life, the only cortical axons capable of conduction are the RAS and C-R axons in layer I.
- Both RAS and C-R axons undergo rapid maturation from the prenatal period to around 6 months of age.
- RAS and C-R axons form a very nonspecific projection, with each axon running for millimeters and contacting a large number of dendrites, producing top-down stimulation of neurons in the deeper cortical layers.
- fMRI confirms that layer I activates all areas of the auditory cortex, with adult-like activation by 3 months of age.
- During the latter half of the first year of life, axons resembling primate MGm projections reach layer I.
- The P2 and derived MMR are both present at birth and mature rapidly after birth.
- The time constants of maturation of P2 and MMR are the same as that of ABR wave V.
- Age 6 months is an anatomical turning point, the time when the C-R system disappears, RAS input appears reduced, and thalamic input to layer I arrives.
- Age 6 months (4–8 months) is a physiological turning point at which the MMR turns from positive to negative, and P2 wave latency begins to shorten to its adult value.
- Age 6 months is a behavioral turning point, the time when, for example, universal discrimination of language sounds begins to regress.
- The layer I arousal system remains functional across the life span, though reduced compared to neonates.
- The MMN and P2 are robustly present in late-onset (postlingual) deafness, probably because they mature very early.

5.2 Anatomical and Physiological Correlates of Perception

Perception is the linking of a stimulus to its meaning or significance. From the anatomical point of view, it is a function of the cortex, driven initially by bottom-up stimulation from the thalamus and continuing with transcortical input into superficial layers. Both perception and cortical connections begin to develop slowly in the second half of the first year of life, and continue into late childhood/teen/adult years. This phase is characterized by thalamocortical input; it forms the basis of perception and the acquisition of language.

5.2.1 Deep Cortical Layers and Pa–Nb/P1

- Filament maturation in the acoustic radiation, the proximal part of the pathway from thalamus to auditory cortex, begins shortly after birth.

- Myelination of the axons in the acoustic radiation begins at around 4 months postnatal and continues to about 4 years.
- Because of its straight medial-to-lateral course, the acoustic radiation will not generate a far-field potential.
- Filament maturation of the distal thalamic axons into the deeper cortical layers is an extended process, occurring progressively from the later part of the first year of life to age 5–6.
- The onset of thalamocortical input into the cortex from about 6 months of age coincides with the regression of layer I and the decline in universal speech sound discrimination.
- The Pa–Nb complex is not present at birth but is mature by age 5. It may result from excitatory (Pa) and inhibitory (Nb) PSPs in layer IV.
- Pa–Nb is unaffected by late-onset deafness, possibly because the thalamic input to layer IV is relatively mature by age 2–3 years.
- P1 is likely generated by cortical deep layers activated by thalamocortical axons, and thus should reflect perceptual processing.
- Because P1 is a deep-layer cortical potential, it should not appear until after 6 months of age. However, the evoked potential component usually labeled P1 is a mix of the middle latency component Pb and the “real” P1, which results from the output of the cholinergic PPT.
- P1 is well developed by age 5–6, which is a good fit with the histology.
- As shown by deprivation studies, maturation of P1 is dependent on activity, measured as time-in-sound. This implies that myelination is a factor, but it may also reflect synaptic maturation, as P1 is generated by a dipole field that requires synchronous, fast-responding dendritic synapses in layer III/IV with return current in the apical dendrite.

5.2.2 Superficial Cortical Layers and N1

- Histological studies show that input into cortical layers II–III matures at age between 6 and 12 years.
- All input to layers II–III is intracortical, with no extrinsic sources (ignoring monoamine networks). Thus, maturation of layer II and III axons significantly broadens the scope of intracortical processing, both within and between hemispheres, and provides a basis for changes in perception in later childhood and beyond.
- Excitatory synapses in the upper layers generate the negative scalp-recorded wave (sink) N1.
- N1 matures gradually from age 6 to 12.
- The N1 wave partly overlaps P1, and thus partly masks it, accounting for the P1 decrease in latency past 6 years of age.
- Absence of N1 after deafness during early years implies an impact of deprivation on maturation of intracortical deep-to-superficial layer axons.

6 Final Summary

This chapter has gone further than any other paper in correlating and explaining human anatomical, histological, electrophysiological, and functional changes during the long maturation process. What is more important, several apparent inconsistencies between structure and early human auditory function, particularly with respect to the mismatch responses, have been cleared up.

The general peripheral to central gradient of auditory system maturation is visible in the successive maturation of the ABR waves I–V and the later maturation of some of the obligatory AEP components. However, histological evidence abounds that the brain stem matures in unison as far as axon neurofilament and myelination are concerned. Thus the gradient is largely the result of activity and synchrony dependent synaptic maturation. Maturation time constants increase stepwise from the periphery (4 weeks), via brain stem (6 months) to the thalamo-cortical system (6 years), as shown by histology, structural imaging, electrophysiology, and functional imaging, and correspond well with behavioral indices of sensory discrimination and perception. What distinguishes humans from other mammals is the very long time constant for thalamocortical maturation, whereas the maturation of the auditory periphery and brain stem follows largely the same time course.

Three pathways to the auditory cortex can be distinguished and appear to mature along very different timelines. The ones that mature early—the reticular activating system pathway and the extralemnisal, nontopographically organized, pathway—are generally adult-like at the end of the maturation of the neural discrimination system, that is, by 1.5–2 years of age. The lemniscal, tonotopically organized, pathway appears the slowest to mature, well into the late teens or early twenties, and correlates with the maturation of the perceptual system.

7 Appendix: Modeling Maturation Changes

Maturation follows a time course to adult values that can be modeled by a sum of exponential functions (Eggermont 1988). Each exponential can be characterized by a rate of change, the time constant, and by a time to reach maturity, that is, the adult value. Each exponential is assumed to describe a single maturing mechanism or more likely a maturing structure. Time constants increase from periphery to cortex, reflecting the hierarchical gradient of maturation.

Models of growth and development as used in biology are characterized by an age (t) dependent rate of change $M(t)$ defined, for our purpose, as the time derivative of latency L (but it could also be relative white matter volume, etc.):

$$M(t) = dL(t) / dt \tag{A.1}$$

The most convenient model in a biological context is $M(t) = r * L(t)$, that is, the rate of change in latency (white matter density, etc.) at any point in time is proportional to the actual value of the latency (white matter density, etc.) at that point in time. If $r > 0$, there is exponential growth with age and if $r < 0$, there is exponential decay. The decay is made explicit in the following by using a negative (-) sign in front of the r . Here, r is the so-called maturation rate constant, and its inverse is called the maturation time constant $\tau = 1/r$.

Realizing that $M(t) = dL(t) / dt = rL(t)$ implies that:

$$L(t) = L(0) * \exp(-rt) + L(\text{adult}) \quad (\text{A.2})$$

where $L(0)$ represents the latency value at time $t=0$. Plotting the natural log of the latency difference with adult value, $\ln[L(t) - L(\text{adult})]$, as a function of age will result in a straight line curve fit with slope equal to $-r$. This r -value is a true characteristic of the maturation process; the latency drops to $1/e$ ($e=2.72$; the base of the natural logarithm) of its value for every τ units of age.

The decaying exponential model for latency changes with age can easily be extended to the case where more than one process is producing changes, presuming that these processes are independent. For instance, one could interpret (Eggermont 1988) the latency changes in ABR wave V as reflecting the latency changes in wave I (peripheral changes), and those in the central conduction time (the I-V delay). A model that takes this concurrent maturation into account is one that shows two maturation rate constants:

$$dL_V(t) / dt = r_I L_I(t) + r_V L_V(t) \quad (\text{A.3})$$

where $L_V(t)$ is the latency of wave V at age t . So waves I and V have their own maturation rates, and this gives rise to a sum of exponentials model:

$$L_V(t) = L_V(\text{adult}) + L_I(0) * \exp(-r_I t) + L_V(0) * \exp(-r_V t) \quad (\text{A.4})$$

In case of abnormal maturation, it is important to know if the onset of maturation of certain processes is delayed, whether the maturation rates are affected, or that both onset and maturation rate are affected. Children fitted with cochlear implants some time after they became deaf form an example population. In this case, maturation is in fact arrested at the onset of deafness and is restarted at the time of implant activation (Ponton et al. 1996a, b). The exponential curve fit can easily accommodate both delay (t_d) and rate of maturation change:

$$L(t) = L(\text{adult}) + L(0) * \exp[-r(t - t_d)] \quad (\text{A.5})$$

Here, t_d is the duration of the deafness, and $t - t_d$ characterizes the hearing age or time-in-sound. If neither $L(\text{adult})$ nor r is affected, the maturation should follow the same time course as in normal development if the time axis is hearing age, which for normal hearing subjects is equal to the chronological age.

An extension of this model is warranted when dealing with increasing values, for example, white matter volume, head size, and so forth. Typically, this growth cannot exceed a certain maximum value, and a modification of the model for limited growth can be as follows:

$$dA(t) / dt = rA(t) \{A(\text{adult}) - A(t)\} / A(\text{adult}) = rA(t) \{1 - A(t) / A(\text{adult})\}. \quad (\text{A.6})$$

Alternatively, one could interpret $r(t) = r \{1 - A(t) / A(\text{adult})\}$ as an age-dependent maturation rate “constant.” As long as $A(t)$ is small compared to $A(\text{adult})$, $r(t) = r$ and there is exponential growth. When $A(t)$ gets bigger, the maturation rate drops. This model shows a sigmoidal dependence of $A(t)$:

$$A(t) = A(\text{adult}) / \left\{ 1 + \left\{ Y \leq (A(\text{adult}) - N(0)) / f / N(0) \right\} \exp(-rt) \right\} \quad (\text{A.7})$$

It is evident that for large t values where $\exp(-rt)$ approaches 0, $A(t)$ approaches adult values.

Acknowledgments This work was supported by the Alberta Heritage Foundation for Medical Research, and by the Campbell McLaurin Chair for Hearing Deficiencies.

References

- Aitkin, L. M., Kudo, M., & Irvine, D. R. F. (1988). Connections of the primary auditory cortex in the common marmoset, *Calithrix jacchus jacchus*. *Journal of Comparative Neurology*, 269, 235–248.
- Albrecht, R., Suchodoletz, W., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clinical Neurophysiology*, 111, 2268–2276.
- Anderson, A. W., Marois, R., Colson, E. R., Peterson, B. S., Duncan, C. C., Ehrenkranz, R. A., Schneider, K. C., Gore, J. C., & Ment, L. R. (2001). Neonatal auditory activation detected by functional magnetic resonance imaging. *Magnetic Resonance Imaging*, 19, 1–5.
- Barnet, A. B., Ohlrich, E. S., Weiss, I. P., & Shanks, B. (1975). Auditory evoked potentials during sleep in normal children from ten days to three years of age. *Electroencephalography and Clinical Neurophysiology*, 39, 29–41.
- Beaulieu, C., & Colonnier, M. (1985). A laminar analysis of the number of round-asymmetrical and flat-symmetrical synapses on spines, dendritic trunks, and cell bodies in area 17 of the cat. *Journal of Comparative Neurology*, 231, 180–189.
- Bishop, D. V., Hardiman, M., Uwer, R., & von Suchodoletz, W. (2007). Maturation of the long-latency auditory ERP: Step function changes at start and end of adolescence. *Developmental Science*, 10, 565–575.
- Brodmann, K. (1908). Beitrage zuer histologischen Lokalisation der Grosshirnrinde: VI Mitteilung: Die Cortexgliederung der Menschen. *Journal of Psychiatry and Neurology*, 10, 231–246.
- Bruneau, N., Roux, S., Guerin, P., Barthelemy, C., & Lelord, G. (1997). Temporal prominence of auditory evoked potentials (N1 wave) in 4–8-year-old children. *Psychophysiology*, 34, 32–38.
- Bürgel, U., Amunts, K., Hoemke, L., Mohlberg, H., Gilsbach, J. M., & Zilles, K. (2006). White matter fiber tracts of the human brain: Three-dimensional mapping at microscopic resolution, topography and intersubject variability. *NeuroImage*, 29, 1092–1105.

- Burton, H., & Jones, E. G. (1975). The posterior thalamic region and its cortical projection in New World and Old World monkeys. *Journal of Comparative Neurology*, 168, 249–302.
- Caulier, L. J., & Connors, B. W. (1994). Synaptic physiology of horizontal afferents to layer I in slices of rat SI neocortex. *The Journal of Neuroscience*, 14, 751–762.
- Ceponienė, R., Rinne, T., & Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clinical Neurophysiology*, 113, 870–882.
- Cone, N. E., Burman, D. D., Bitan, T., Bolger, D. J., & Booth, J. R. (2008). Developmental changes in brain regions involved in phonological and orthographic processing during spoken language processing. *NeuroImage*, 41, 623–635.
- Courchesne, E., Chisum, H. J., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hinds, S., & Press, G. A. (2000). Normal brain development and aging: Quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology*, 216, 672–682.
- Dehaene-Lambertz, G., & Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience*, 16, 1375–1387.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298, 2013–2015.
- de la Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2006). Thalamic connection of the auditory cortex in marmoset monkeys: Core and medial belt regions. *Journal of Comparative Neurology*, 496, 72–96.
- del Rio, J. A., Martinez, A., Fonseca, M., Auladell, C., & Soriano, E. (1995). Glutamate-like immunoreactivity and fate of Cajal-Retzius cells in the murine cortex as identified by calretinin antibody. *Cerebral Cortex*, 1, 13–21.
- Devous, M. D. Sr., Altuna, D., Furl, N., Cooper, W., Gabbert, G., Ngai, W. T., Chiu, S., Scott, J. M. 3rd, Harris, T. S., Payne, J. K., & Tobey, E. A. (2006). Maturation of speech and language functional neuroanatomy in pediatric normal controls. *Journal of Speech Language and Hearing Research*, 49, 856–866.
- Draganova, R., Eswaran, H., Murphy, P., Huutilainen, M., Lowery, C., & Preissl, H. (2005). Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *NeuroImage*, 28, 354–361.
- Draganova, R., Eswaran, H., Murphy, P., Lowery, C., & Preissl, H. (2007). Serial magnetoencephalographic study of fetal and newborn auditory discriminative evoked responses. *Early Human Development*, 83, 199–207.
- Eggermont, J. J. (1988). On the rate of maturation of sensory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 70, 293–305.
- Eggermont, J. J. (2006). Electric and magnetic fields of synchronous neural activity propagated to the surface of the head: Peripheral and central origins of AEPs. In R. R. Burkard, M. Don, & J. J. Eggermont (Eds.), *Auditory evoked potentials* (pp. 2–21), Baltimore: Lippincott Williams & Wilkins.
- Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: From single-units to evoked potentials. *Audiology & Neuro-Otology*, 7, 71–99.
- Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: Correlations with changes in structure and speech perception. *Acta Oto-Laryngologica*, 123, 249–252.
- Eggermont, J. J., & Salamy, A. (1988). Maturation time course for the ABR in preterm and full term infants. *Hearing Research*, 33, 35–47.
- Eggermont, J. J., Ponton, C. W., Coupland, S. G., & Winkelaar, R. (1991). Maturation of the traveling-wave delay in the human cochlea. *Journal of the Acoustical Society of America*, 90, 288–298.
- Eggermont, J. J., Brown, D. K., Ponton, C. W., & Kimberley, B. P. (1996). Comparison of distortion product otoacoustic emission (DPOAE) and auditory brain stem response (ABR) traveling wave delay measurements suggests frequency-specific synapse maturation. *Ear & Hearing*, 17, 386–394.
- Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as function of age. *Journal of the Acoustical Society of America*, 107, 2704–2710.

- Elliott, L. L. (1979). Performance of children aged 9–17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *Journal of the Acoustical Society of America*, 66, 651–653.
- Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport*, 13, 1251–1254.
- Galaburda, A., & Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *Journal of Comparative Neurology*, 190, 597–610.
- Gilley, P. M., Sharma, A., Dorman, M., & Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clinical Neurophysiology*, 116, 648–657.
- Gilley, P. M., Sharma, A., & Dorman, M. F. (2008). Cortical reorganization in children with cochlear implants. *Brain Research*, 1239, 56–65.
- Gomes, H., Dunn, M., Ritter, W., Kurtzberg, D., Brattson, A., Kreuzer, J. A., & Vaughan, H. G. Jr. (2001). Spatiotemporal maturation of the central and lateral N1 components to tones. *Developmental Brain Research*, 129, 147–155.
- Gomot, M., Giard, M. H., Roux, S., Barthélémy, C., & Bruneau, N. (2000). Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *NeuroReport*, 11, 3109–3012.
- Gordon, K. A., Papsin, B. C., & Harrison, R. V. (2005). Effects of cochlear implant use on the electrically evoked middle latency response in children. *Hearing Research*, 204, 78–89.
- Gordon, K. A., Papsin, B. C., & Harrison, R. V. (2006). An evoked potential study of the developmental time course of the auditory nerve and brainstem in children using cochlear implants. *Audiology & Neuro-Otology*, 11, 7–23.
- Hackett, T. A., Stepniewska, I., & Kaas, J. H. (1998a). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, 394, 475–495.
- Hackett, T. A., Stepniewska, I., & Kaas, J. H. (1998b). Thalamocortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, 400, 271–286.
- Hackett, T. A., Stepniewska, I., & Kaas, J. H. (1999). Callosal connections of the parabelt auditory cortex in macaque monkeys. *European Journal of Neuroscience*, 11, 856–866.
- Hafner, H., Pratt, H., Joachims, Z., Feinsod, M., & Blazer, S. (1991). Development of auditory brainstem evoked potentials in newborn infants: A three-channel Lissajous's trajectory study. *Hearing Research*, 51, 33–47.
- Harrison, J. B., Woolf, N. J., & Buchwald, J. S. (1990). Cholinergic neurons of the feline pontomesencephalon. I. Essential role in 'wave A' generation. *Brain Research*, 520, 43–54.
- Hashikawa, T., Molinari, M., Rausell, E., & Jones, E. G. (1995). Patchy and laminar termination of medial geniculate axons in monkey auditory cortex. *Journal of Comparative Neurology*, 362, 195–208.
- Hashimoto, I., Ishiyama, Y., Yoshimoto, T., & Nemoto, S. (1981). Brain-stem auditory-evoked potentials recorded directly from human brain-stem and thalamus. *Brain*, 104, 841–859.
- He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, 19, 878–892.
- Hestrin, L., & Armstrong, W. E. (1996). Morphology and physiology of cortical neurons in layer I. *The Journal of Neuroscience*, 16, 5290–5300.
- Hoffman, P. N., Griffin, J. W., & Price, D. L. (1984). Control of axonal caliber by neurofilament transport. *Journal of Cell Biology*, 99, 705–714.
- Hüppi, P. S., & Dubois, J. (2006). Diffusion tensor imaging of brain development. *Seminars in Fetal Neonatal Medicine*, 11, 489–497.
- Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, 387, 167–178.
- Imamoto, K., Karasawa, N., Isomura, G., & Nagatsu, I. (1994). Cajal-Retzius neurons identified by GABA immunohistochemistry in layer I of the rat cerebral cortex. *Neuroscience Research*, 20, 101–115.
- Javitt, D. C., Steinschneider, M., Schroeder, C. E., & Arezzo, J. C. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: Implications for schizophrenia. *Proceedings of the National Academy of Sciences of the USA*, 93, 11962–11967.

- Jiang, Z. D., Zheng, M. S., Sun, D. K., & Liu, X. Y. (1991). Brainstem auditory evoked responses from birth to adulthood: Normative data of latency and interval. *Hearing Research*, 54, 67–74.
- Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech Language and Hearing Research*, 43, 144–157.
- Kinney, H. C., Brody, B. A., Kloman, A. S., & Gilles, F. H. (1988). Sequence of central nervous system myelination in human infancy. II. Patterns of myelination in autopsied infants. *Journal of Neuropathology and Experimental Neurology*, 47, 217–234.
- Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews*, 56, 259–269.
- Kraus, N., Smith, D. I., Reed, N. L., Stein, L. K., & Cartee, C. (1985). Auditory middle latency responses in children: Effects of age and diagnostic category. *Electroencephalography and Clinical Neurophysiology*, 62, 343–351.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants 6 months of age. *Science*, 255, 606–608.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, 13, 47–51.
- Langworthy, O. R. (1933). Development of behavioral patterns and myelination of the nervous system in the human fetus and infant. *Contributions to Embryology (Carnegie Institute of Washington)*, 24, 1–57.
- Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *NeuroImage*, 40, 1044–1055.
- Lieberman, A., Sohmer, H., & Szabo, G. (1973). Standard values of amplitude and latency of cochlear audiometry (electro-cochleography). Responses in different age groups. *Archiven Klinische und Experimentelle Ohren Nasen und Kehlkopfheilkunde*, 203, 267–273.
- Luetheke, L. E., Krubitzer, L. A., & Kaas, J. H. (1989). Connections of primary auditory cortex in the New World monkey, *Saguinus*. *Journal of Comparative Neurology*, 285, 487–513.
- Lütkenhöner, B., & Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol & Neuro-Otology*, 3, 191–213.
- Marin-Padilla, M., & Marin-Padilla, T. M. (1982). Origin, prenatal development and structural organization of layer I of the human cerebral (motor) cortex. A Golgi study. *Anatomy and Embryology (Berlin)*, 164, 161–206.
- Meyer, G., & Goffinet, A. M. (1998). Prenatal development of reelin-immunoreactive neurons in the human neocortex. *Journal of Comparative Neurology*, 397, 29–41.
- Meyer, G., & González-Hernández, T. (1993). Developmental changes in layer I of the human neocortex during prenatal life: A DiI-tracing, AChE and NADPH-d histochemistry study. *Journal of Comparative Neurology*, 338, 317–336.
- Michalewski, H. J., Starr, A., Nguyen, T. T., Kong, Y. Y., & Zeng, F. G. (2005). Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. *Clinical Neurophysiology*, 116, 669–680.
- Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. *Physiological Reviews*, 65, 37–100.
- Mochizuki, Y., Go, T., Ohkubo, H., & Motomura, T. (1983). Development of human brainstem auditory evoked potentials and gender differences from infants to young adults. *Progress in Neurobiology*, 20, 273–285.
- Møller, A. R., Jannetta, P. J., & Sekhar, L. N. (1988). Contributions from the auditory nerve to the brain-stem auditory evoked potentials (BAEPs): Results of intracranial recording in man. *Electroencephalography and Clinical Neurophysiology*, 71, 198–211.
- Moore, J. K. (2002). Maturation of human auditory cortex: Implications for speech perception. *Annals of Otolology Rhinology and Laryngology Supplement*, 189, 7–10
- Moore, J. K., & Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2, 297–311.

- Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: A timeline of development. *International Journal of Audiology*, 46, 460–478.
- Moore, J. K., Perazzo, L. M., & Braun, A. (1995). Time course of axonal myelination in the human brainstem auditory pathway. *Hearing Research*, 87, 21–31.
- Moore, J. K., Ponton, C. W., Eggermont, J. J., Wu, B. J., & Huang, J. Q. (1996). Perinatal maturation of the auditory brain stem response: Changes in path length and conduction velocity. *Ear & Hearing*, 17, 411–418.
- Moore, J. K., Guan, Y. L., & Shi, S. R. (1997). Axogenesis in the human fetal auditory system, demonstrated by neurofilament immunohistochemistry. *Anatomy and Embryology (Berlin)*, 195, 15–30.
- Moore, J. K., Guan, Y. L., & Shi, S. R. (1998). MAP2 expression in developing dendrites of human brainstem auditory neurons. *Journal of Chemical Neuroanatomy*, 16, 1–15.
- Mukherjee, P., Miller, J. H., Shimony, J. S., Conturo, T. E., Lee, B. C., Almlí, C. R., & McKinstry, R. C. (2001). Normal brain maturation during childhood: Developmental trends characterized with diffusion-tensor MR imaging. *Radiology*, 221, 349–358.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1–21.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24, 375–425.
- Näätänen, R., Lehtokoski, A., Lenne, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385, 432–434.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, H. G., Jr. (1989). Cortical responses to speech sounds and their formants in normal infants: Maturation sequence and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, 73, 295–305.
- Ohlrich, E. S., Barnett, A. B., Weiss, I. P., & Shanks, B. L. (1978). Auditory evoked potential development in early childhood: A longitudinal study. *Electroencephalography and Clinical Neurophysiology*, 44, 411–423.
- Paetau, R., Ahonen, A., Salonen, O., & Sams, M. (1995). Auditory evoked magnetic fields to tones and pseudowords in healthy children and adults. *Journal of Clinical Neurophysiology*, 12, 177–185.
- Pandya, D. N., & Rosene, D. L. (1993). Laminar termination patterns of thalamic, callosal and association afferents in the primary auditory area of the rhesus monkey. *Experimental Neurology*, 119, 220–234.
- Pang, E. W., & Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: An examination of speech and non-speech stimuli. *Clinical Neurophysiology*, 111, 388–397.
- Pang, E. W., Edmonds, G. E., Desjardins, R., Khan, S. C., Trainor, L. J., & Taylor, M. J. (1998). Mismatch negativity to speech stimuli in 8-month-old infants and adults. *International Journal of Psychophysiology*, 29, 227–236.
- Pasman, J. W., Rotteveel, J. J., de Graaf, R., Maassen, B., & Notermans, S. L. H. (1991). Detectability of auditory response components in preterm infants. *Early Human Development*, 26, 129–141.
- Pasman, J. W., Rotteveel, J. J., Maassen, B., & Visco, Y. M. (1999). The maturation of auditory cortical evoked responses between (preterm) birth and 14 years of age. *European Journal of Paediatric Neurology*, 3, 79–82.
- Paus, T., Collins, D. L., Evans, A. C., Leonard, G., Pike, B., & Zijdenbos, A. (2001). Maturation of white matter in the human brain: A review of magnetic resonance studies. *Brain Research Bulletin*, 54, 255–266.
- Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Archives of Neurology*, 51, 874–887.
- Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Developmental Neuropsychology*, 31, 249–278.
- Ponton, C. W., & Eggermont, J. J. (2001). Of kittens and kids: Altered cortical maturation following profound deafness and cochlear implant use. *Audiology & Neuro-Otology*, 6, 363–380.

- Ponton, C. W., & Eggermont, J. J. (2006). Electrophysiological measures of human auditory system maturation: Relationship with neuroanatomy and behavior. In R. R. Burkard, M. Don, & J. J. Eggermont (Eds), *Auditory evoked potentials* (pp. 385–402), Baltimore: Lippincott Williams & Wilkins.
- Ponton, C. W., Eggermont, J. J., Coupland, S. G., & Winkelaar, R. (1992). Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs). *Journal of the Acoustical Society of America*, 91, 1576–1586.
- Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., Kwong, B., & Masuda, A. (1996a). Auditory system plasticity in children after long periods of complete deafness. *NeuroReport*, 8, 61–65.
- Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., & Masuda, A. (1996b). Maturation of human cortical auditory function: Differences between normal-hearing children and children with cochlear implants. *Ear & Hearing*, 17, 430–437.
- Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000a). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, 111, 220–236.
- Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., Kwong, B., Cunningham, J., & Trautwein, P. (2000b). Maturation of the mismatch negativity: Effects of profound deafness and cochlear implant use. *Audiology & Neuro-Otology*, 5, 167–185.
- Ponton, C. W., Eggermont, J. J., Khosla, D., Kwong, B., & Don, M. (2002). Maturation of human central auditory system activity: Separating auditory evoked potentials by dipole source modeling. *Clinical Neurophysiology*, 113, 407–420.
- Pujol, J., Soriano-Mas, C., Ortiz, H., Sebastián-Gallés, N., Losilla, J. M., & Deus, J. (2006). Myelination of language-related areas in the developing brain. *Neurology*, 66, 339–343.
- Rauschecker, J. P., Tian, B., Pons, T., & Mishkin, M. (1997). Serial and parallel processing in rhesus monkey auditory cortex. *Journal of Comparative Neurology*, 382, 89–103.
- Ramon y Cajal, S. (1900). Studies on the human cerebral cortex III: Structure of the acoustic cortex. *Revista Trimestral Micrografica*, 5, 129–183
- Rotteveel, J. J., Colon, E. J., Notermans, L. H., Stoelinga, G. B. A., & Visco, Y. M. (1985). The central auditory conduction at term date and three months after birth. I. Composite group averages of brainstem ABR, middle latency (MLR) and auditory cortical responses (ACR). *Scandinavian Audiology*, 14, 179–186.
- Rotteveel, J. J., Stegeman, D. F., de Graaf, R., Colon, E. J., & Visco, Y. M. (1987). The maturation of the central auditory conduction in preterm infants until three months post term. III. The middle latency auditory evoked response (MLR). *Hearing Research*, 27, 245–256.
- Rubel, E. W., Lippe, W. R., & Ryals, B. M. (1984). Development of the place principle. *Annals of Otolaryngology and Laryngology*, 93, 609–615.
- Sano, M., Kaga, K., Kuan, C. C., Ino, K., & Mima, K. (2007). Early myelination patterns in the brainstem auditory nuclei and pathway: MRI evaluation study. *International Journal of Pediatric Otorhinolaryngology*, 71, 1105–1115.
- Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear & Hearing*, 23, 532–539.
- Sharma, A., Gilley, P. M., Dorman, M. F., & Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *International Journal of Audiology*, 46, 494–499.
- Skinner, J. E., & Yingling, C. D. (1976). Regulation of slow potential shifts in nucleus reticularis thalami by the mesencephalic reticular formation and the frontal granular cortex. *Electroencephalography and Clinical Neurophysiology*, 40, 288–296.
- Spreafico, R., Arcelli, P., Frassoni, C., Canetti, P., Giaccone, G., et al. (1999). Development of layer I of the human cerebral cortex after midgestation: Architectonic findings, immunocytochemical identification of neurons and glia, and in situ labeling of apoptotic cells. *Journal of Comparative Neurology*, 410, 126–142.

- Starr, A., Amlie, R. N., Martin, W. H., & Sanders, S. (1977). Development of auditory function in newborn infants revealed by auditory brainstem potentials. *Pediatrics*, 60, 831–839.
- Stegeman, D. F., Van Oosterom, A., & Colon, E. J. (1987). Far-field evoked potential components induced by a propagating generator: Computational evidence. *Electroencephalography and Clinical Neurophysiology*, 67, 176–187.
- Su, P., Kuan, C. C., Kaga, K., Sano, M., & Mima, K. (2008). Myelination progression in language-correlated regions in brain of normal children determined by quantitative MRI assessment. *International Journal of Pediatric Otorhinolaryngology*, 72, 1751–1763.
- Thai-Van, H., Cozma, S., Boutitie, F., Disant, F., Truy, E., & Collet, L. (2007). The pattern of auditory brainstem response wave V maturation in cochlear-implanted children. *Clinical Neurophysiology*, 118, 676–689.
- Tonnquist-Uhlen, I., Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2003). Maturation of human central auditory system activity: The T-complex. *Clinical Neurophysiology*, 114, 685–701.
- Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology*, 51, 5–15.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 47, 466–472.
- Trojanowski, J. Q., & Jacobson, S. (1975). A combined horseradish peroxidase autoradiographic investigation of reciprocal connections between superior temporal gyrus and pulvinar in squirrel monkey. *Brain Research*, 85, 347–353.
- Weisshaar, B., Doll, T., & Matus, A. (1992). Reorganisation of the microtubular cyoskeleton by embryonic microtubule-associated protein 2 (MAP2c). *Development*, 116, 1151–1161.
- Weitzman, W. D., & Graziani, L. J. (1968). Maturation and topography of the auditory evoked response of the prematurely born infant. *Developmental Psychobiology*, 1, 79–89.
- Werker, J. F., & Tees, R. S. (1984). Cross language speech perception: Evidence for perceptual organization during the first year of life. *Infant Behavioral Development*, 7, 49–63.
- Winer, J. A., & Larue, D. T. (1989). Populations of GABAergic neurons and axons in layer I of rat auditory cortex. *Neuroscience*, 33, 499–515.
- Xu, Z., Marszalek, J. R., Lee, M. K., Wong, P. C., Folmer, J., Crawford, T. O., Hsieh, S. T., Griffin, J. W., & Cleveland, D. W. (1996). Subunit composition of neurofilaments specifies axonal diameter. *Journal of Cell Biology*, 133, 1061–1069.
- Yakolev, P. L., & Lecours, A. R. (1967). The myelogenetic cycles of regional maturation of the brain. In A Minkowski (Ed.), *Regional development of the brain in early life* (pp. 3–70). Oxford: Blackwell.
- Yingling, C. D., & Skinner, J. E. (1976). Selective regulation of thalamic sensory relay nuclei by nucleus reticularis thalami. *Electroencephalography and Clinical Neurophysiology*, 41, 476–482.
- Zecević, N., Milosevic, A., Rakic, P., & Marin-Padilla, M. (1999). Early development and composition of the human primordial plexiform layer: An immunohistochemical study. *Journal of Comparative Neurology*, 412, 241–254.
- Zhang, J., Evans, A., Hermoye, L., Lee, S. K., Wakana, S., Zhang, W., Donohue, P., Miller, M. I., Huang, H., Wang, X., van Zijl, P. C., & Mori, S. (2007). Evidence of slow maturation of the superior longitudinal fasciculus in early childhood by diffusion tensor imaging. *NeuroImage*, 38, 239–247.

Chapter 4

Development of Auditory Coding as Reflected in Psychophysical Performance

Emily Buss, Joseph W. Hall III, and John H. Grose

1 Introduction

Behavioral response to sound can be consistently elicited by 28 weeks of gestational age (Birnholz and Benacerraf 1983), and yet performance on some auditory tasks continues to mature into early adolescence even for relatively simple stimuli (Fior 1972; Maxon and Hochberg 1982; Fischer and Hartnegg 2004). This protracted auditory development could be affected by a number of factors, including anatomical, neural, and cognitive maturation. Some of these factors are discussed by Abdala and Keefe (Chap. 2) and Eggermont and Moore (Chap. 3). Many studies of infants and children have documented developmental effects on basic psychoacoustical tasks, such as stimulus detection or the ability to discriminate changes in frequency or intensity, but little is known about the underlying sources of immaturity.

2 Measurement Techniques

The testing methods that are typically used to characterize auditory perception differ substantially as a function of age. Development of testing methods for young listeners has been guided by the goal of obtaining accurate, reliable, and unbiased measurements within a reasonable amount of time, achieved using procedures that promote sustained interest and motivation on the part of the listener.

E. Buss (✉) • J.W. Hall III • J.H. Grose
Department of Otolaryngology/Head and Neck Surgery,
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
e-mail: ebuss@med.unc.edu; jwh@med.unc.edu; jhg@med.unc.edu

2.1 *Conditioned Reflexes*

Using a habituation/dishabituation paradigm, Shahidullah and Hepper (1994) showed that by 35 weeks gestational age, fetuses can distinguish between a 250-Hz tone and a 500-Hz tone. In that study, stimuli were delivered to a modified head-*phone* positioned on the mother's abdomen, near the head of the fetus. Tones were presented for 2 s, with one tone every 7 s. Stimuli were presented at approximately 110 dB(A), and these stimuli were reported to be inaudible to the mother. The fetus's activity level was monitored by ultrasound. The experimenter continued presenting the first sound until the fetus was still, indicating habituation. The second sound was then presented, and an increase in fetal activity was judged to reflect dishabituation, and hence ability to discriminate the two sounds. Similar methods have been used to assess perception in young infants.

One limitation of this procedure in general is that inferring sensitivity from conditioned reflexes relies on the innate tendency to respond to novel stimuli. Repeated presentations reduce novelty, and so reduce the probability that the infant will respond even if the stimulus change is perceived. Several authors have argued that that visual reinforcement of response to sound produces more reliable threshold estimates over time than observation methods (Moore et al. 1977; Trehub et al. 1981; Primus 1991). Therefore, the techniques available to measure fetal hearing may underestimate sensitivity.

2.2 *Observer-Based Methods*

The observer-based psychophysical method (Olsho et al. 1987b) is widely used to obtain psychoacoustical measurements in infants. In this method, the infant is seated on a caregiver's lap inside a sound booth. An experimenter inside the booth uses quiet toys to focus the infant's attention. When the infant is quiet and facing the midline an experimenter outside the booth (the observer) initiates a stimulus trial. The observer then monitors the behavior of the infant and tries to determine whether or not the infant heard a change in the auditory stimulus. The observer can base judgments on any of an open set of criteria (e.g., eye-widening, brow movement, etc.). When the infant's behavior allows the observer to identify a signal trial correctly, a Plexiglas box inside the booth illuminates, revealing a toy that performs a brief animation. Before data collection there is a training phase with suprathreshold stimuli to give the observer an opportunity to determine reliable features of the infant's behavioral response to stimulus change and to demonstrate acceptable stimulus control. Observer-based audiometry is similar to visual reinforcement audiometry (VRA; Moore et al. 1975; Primus 1991; Gravel and Traquina 1992), with the primary difference being that VRA is based on the infant making a head-turn toward the source of the sound, whereas the observer-based method uses any available behavioral response. This distinction is important because infants younger than about 5 months cannot be conditioned to make reliable head-turn responses (Moore et al. 1977).

Results obtained with observer-based methods generally agree with VRA data in 6-month-old infants (Olsho et al. 1987b). Observer-based methods can be used to determine threshold with infants as young as 2–4 months of age but appear to be unreliable in neonates (Hicks et al. 2000; Tharpe and Ashmead 2001).

One strength of the observer-based method is that the resulting data are relatively free of observer bias. The observer is blind to whether or not a signal interval occurred, and so must rely solely on the infant's behavior. Further, both the care giver and the experimenter inside the booth wear headphones playing masking noise, eliminating the possibility that they might inadvertently influence the infant's response. Most infant studies have employed up–down adaptive staircase methods (Taylor and Creelman 1967; Levitt 1971) or fixed-block testing, in which sensitivity is measured at a range of stimulus values. Observer-based testing has been applied successfully to a wide range of auditory tasks, including detection in quiet (Olsho et al. 1988; Werner and Boike 2001), masked detection (Werner 1999; Werner and Boike 2001), gap detection (Werner et al. 1992), and frequency discrimination (Olsho et al. 1987a).

One feature of the observer-based method that differs from traditional psychophysical methods used with adults is that the listening trial is temporally uncertain in most cases. Because the observer initiates the trial when the infant appears attentive, there is a great deal of variability in the intertrial interval. A recent study by Werner et al. (2009) tested the hypothesis that the poor thresholds often reported in infant psychophysical studies could be related to temporal uncertainty. In that study an auditory cue preceding the listening interval improved sensitivity of adult listeners but failed to improve performance for 7- to 9-month-olds. A second experiment showed that, whereas infants' performance did not improve with introduction of the cue, introducing an inaccurate cue hurt performance under some conditions. These results were interpreted as showing that infants do develop temporal expectancies, but that auditory cues themselves can be distracting. Whereas the 7- to 9-month-olds in the Werner et al. (2009) study did not seem to benefit from pretrial cues, previous VRA data reported by Primus (1988) indicate that infants 11–15 months of age may benefit from such cues, though cues were not presented before catch trials and infant/adult detection differences were found even with the cue in that experiment. This result indicates that older infants are able to benefit from cues that reduce temporal uncertainty. Although temporal uncertainty could elevate thresholds in the observer-based threshold estimation paradigm used to assess hearing in infancy, this is a relatively small effect compared to the threshold differences obtained in many studies of infant hearing.

Although the observer-based method has been most widely used in the testing of infants, several studies have used this technique to assess hearing in young children. For example, Grieco-Calub et al. (2008) investigated localization ability in young children with bilateral cochlear implants and demonstrated that observer-based psychophysical procedures can be effective in testing children as old as 3 years of age. Analogous VRA methods have been used to assess hearing in children as old as 3–4 years of age (Schneider et al. 1986).

2.3 *Forced-Choice Procedures*

Psychoacoustical testing of children 4 years of age and older is usually conducted using some type of forced-choice procedure in which the child presses a button or points to a picture that is associated with a listening interval corresponding to signal presentation. Although some studies employing these procedures have tested children as young as 3 years of age (Allen et al. 1989; Wightman et al. 1989), the youngest age at which children are tested in many studies using these methods is 4 or 5 years. These studies typically employ some type of visual reinforcement, such as brief animations (Wightman et al. 1989) or progress in the completion of a jigsaw puzzle image throughout a block of trials (Allen and Wightman 1994). The three-alternative forced choice method is common because it requires minimal instruction; the child can be instructed to select the interval with the “different” sound, eliminating the requirement that children understand and remember qualitative descriptions of the target stimulus (e.g., “higher pitch”). Some studies have also used suprathreshold reminder trials to help the listener maintain attention on the relevant features of the stimulus (Litovsky 2005).

Most studies of school-age children have used up–down adaptive staircase methods (Taylor and Creelman 1967; Levitt 1971) wherein task difficulty is adaptively varied based on the pattern of recent observer responses. Using this method the signal level at the outset of a track is suprathreshold, and the task becomes gradually more difficult as the track approaches and then hovers around threshold. Another adaptive method that has been used to assess sensitivity is the maximum-likelihood method, as described by Green (1993). This procedure starts with a family of predefined psychometric functions and a target percent correct. After each signal presentation the function that best fits all responses obtained up to that point in the track is identified, and the signal level on the next trial is determined based on the target percent correct on that function. In this procedure, task difficulty typically changes dramatically from trial to trial, with signal levels below threshold frequently presented in the first few trials of a threshold estimation track. In contrast to adaptive procedures, for the method of constant stimuli the signal level on a given trial is unrelated to the listener’s responses, being based instead on a random ordering of predetermined signal levels.

These procedures for measuring thresholds in children appear to result in similar threshold estimates. Buss et al. (2001) tested a group of 6- to 11-year-old children and compared threshold estimates obtained with an up–down adaptive staircase, maximum likelihood procedures, and a method of constant stimuli. The study was motivated in part by the hypothesis that listening experience provided by the progressively more difficult signal detection in the staircase method might result in better, more reliable performance in children. Data failed to support this hypothesis; all three methods produced similar estimates of threshold for both simultaneous and backward masking. It was concluded that an “orderly” progression of task difficulty was not of critical importance in thresholds for school-age children. Although the adaptive methods that converge upon a desired percent correct point are very time efficient, methods that estimate percent correct for a range of stimulus values can be of great value in experiments in which it is important to ascertain the slope of the psychometric functions (Allen and Wightman 1994; Buss et al. 2009), as discussed in more detail below.

Whereas psychometric function slope can often be estimated based on data from adaptive tracks, sequential dependencies and concentration of trials at one point on the function (“threshold”) introduce inaccuracies (Leek et al. 1992).

3 Auditory Sensitivity in Quiet

Detection thresholds in quiet improve with increasing age from infancy through the early school-age years (Schneider et al. 1986; Olsho et al. 1988; Trehub et al. 1988). Several studies on this issue indicate gradual improvement of threshold through development, with an approximately 15–25 dB difference between adults and infants narrowing to approximately 5–10 dB by age 5 years. Detection thresholds improve particularly rapidly between 3 and 6 months of age (Tharpe and Ashmead 2001). Among the studies that have examined the effect of signal frequency, some have found more adult-like thresholds at higher frequencies (e.g., 4,000 Hz) than at lower frequencies (e.g., 500 Hz), a trend that has been reported in infants 6 months of age and older and in school-age children (Schneider et al. 1986; Olsho et al. 1988; Trehub et al. 1988). Although the conductive properties of the auditory periphery continue to mature well into childhood, resulting in age-related differences in the transmission of sound into the cochlea (see Werner, Chap. 1), this is unlikely to account for the time course of observed threshold differences. In addition, evidence from electrophysiological and otoacoustic emissions studies suggests that the developmental differences in detection threshold that persist into the early school years are not likely to arise from cochlear or neural factors associated with the sensitivity of the auditory periphery (Folsom and Wynne 1987; Bargones and Burns 1988; Abdala and Folsom 1995). Although there is no definitive explanation for the poor hearing thresholds in quiet demonstrated by young listeners, some possible sources that have been suggested include reduced motivation, attention, memory, or cognitive ability in young listeners; a higher level of internal noise as manifested by increased variability in the neural representation of the stimulus; and higher levels of self-generated acoustical noise (e.g., respiration, heart beat, movement) present at the level of the cochlea (Nozza and Wilson 1984; Schneider et al. 1986; Buss et al. 2009). These accounts are revisited in more detail below.

4 Frequency, Intensity, and Loudness

4.1 Frequency Discrimination

Frequency discrimination is related to pitch perception, a topic that is covered by Trainor and Unrau, Chap. 8. There is a marked change in frequency discrimination abilities between 3 and 6 months of age. At 3 months of age, frequency discrimination is poorer at 4,000 Hz than at 500 Hz, but this pattern reverses by 6 months of age, at which point performance is better at 4,000 Hz than at 500 Hz (Olsho et al. 1987a).

By 6–12 months of age, frequency discrimination is nearly adult-like at 4,000 Hz, with mean thresholds of about 2%, compared to adult thresholds of just under 1% in comparable stimulus and test conditions. However, thresholds at 500 and 1,000 Hz are poorer than adults', with mean thresholds of about 3% for 6- to 12-month-old infants at these frequencies (Olsho et al. 1987a). Comparable results were reported by Aslin (1989). In that study, 6- to 9-month-old infants were able to discriminate the direction of a frequency sweep stimulus for frequency changes of about 3–4%.

Despite this relatively good frequency discrimination performance in studies of infants, frequency discrimination in children is often reported to be relatively poor. Jensen and Neff (1993b) measured frequency discrimination for a 440-Hz pure tone. Thresholds improved between 4 and 6 years of age, from approximately 14% to approximately 1.5%, with the oldest children performing on average slightly more poorly than adults. It was noted that the developmental trends in these data were more pronounced than for intensity discrimination under comparable conditions. Some studies report being unable to measure frequency difference limens reliably in young school-age children, with a third or more 5- to 6-year-olds disqualified because of failure to learn the task or inconsistent data (Thompson et al. 1999; Halliday et al. 2008).

Interestingly, the details of the frequency discrimination task appear to be more critical to the developmental effects observed than those for other tasks, such as pure tone detection in quiet. For example, Olsho et al. (1987b) report better frequency discrimination performance in infant listeners for fixed block than adaptive testing procedures. Moore et al. (2008) proposed that frequency discrimination is more difficult for children than other tasks because of demands on attention and memory. It is therefore likely that many of the current procedures used to evaluate frequency discrimination in children are limited by factors other than the fidelity of the representation of sound in the auditory system. This idea received some support from Sutcliffe and Bishop (2005), who argued that the choice of psychophysical method has a large impact on frequency discrimination in young children, perhaps because of temporal focus of attention. Indeed, differences in testing procedures could be responsible for the good frequency discrimination performance in infants relative to children. Infant studies of frequency discrimination have traditionally used trains of pulsed stimuli, with the signal interval characterized by a change in frequency of several sequential bursts (Olsho et al. 1987a, b; Aslin 1989). This streaming presentation mode could reduce memory load on the listener and facilitate performance compared to the gated presentation typically used in studies with children. The possible role of memory in frequency discrimination data is bolstered by the finding that auditory memory for pitch develops between 6 and 10 years of age, as shown by adjusting the duration of the interstimulus interval (Keller and Cowan 1994).

4.2 Intensity Discrimination

Intensity discrimination is the ability to detect a change in the presentation level of a sound, a feature that could be related to neural firing rate, phase locking, or recruitment of auditory fibers neighboring the fibers tuned to the signal frequency, particularly

for narrowband stimuli. Intensity discrimination is assessed by measuring the ability to detect a change in level of an ongoing stimulus or a difference in level across multiple presentations of a gated stimulus. Developmental data on intensity discrimination results are somewhat mixed, due perhaps in part to these methodological differences in estimating thresholds.

Sinnott and Aslin (1985) presented a pulsed 1,000-Hz tone, and the ability to detect a change in the standard (60-dB SPL) presentation level was estimated for adults and 7- to 9-month-old infants. The mean threshold for detecting an increase in stimulus level was 6.2 dB for infants and 1.8 dB for adults. In contrast, infants were unable to detect a reduction in stimulus levels at the limits of the procedure, whereas adults obtained a threshold of 1.4 dB. Failure of infants to detect a reduction in stimulus level was attributed to the “unnatural” aspects of decrement detection, perhaps related to the trend for important features of speech to receive stress through increased intensity. Infants’ increment detection thresholds were comparable to those reported by Bull et al. (1984) for intensity discrimination with synthesized multisyllabic speech stimuli in 5- to 11-month-olds. It was argued that these results reflect the ability to detect linguistically relevant changes in intensity in infancy.

Berg and Boswell (1998) measured the ability to detect an increase in level of a continuous, 2-octave band of noise in 7-month-olds and adult listeners. They found that infants were poorer than adults at detecting the increment, but that this age effect was larger for a stimulus center frequency of 400 Hz than 4,000 Hz. The age effect was also larger for 10-ms increments than for 100-ms increments. This pattern of results is similar to that noted for detection thresholds, where the largest age effect was observed for brief low-frequency stimuli. Using stimuli and procedures very similar to those of the 1998 study, Berg and Boswell (2000) measured the ability to detect an increase in level of a continuous, 2-octave band of noise in 1- to 3-year-olds and adults. Thresholds were higher for children than adults when the standard level was 35 dB SPL, but increasing the level of the standard improved performance and reduced the adult/child difference. Thresholds were comparable for the 3-year-old children and adults when the standard was presented at 55 dB SPL, a pattern of results that was similar at 400- and 4,000-Hz center frequencies. This series of experiments is consistent with a relatively rapid, frequency-specific development in the perception of intensity, with adult-like intensity discrimination by early preschool-age for some conditions.

Other studies have found evidence of a maturation of intensity discrimination abilities that extends into school-age years. Jensen and Neff (1993b) measured intensity discrimination in 4- to 6-year-old children and adults for a 70-dB SPL, 440-Hz pure tone. They report that thresholds were adult-like for some 4-year-olds and most 5- to 6-year-olds. In contrast, Buss et al. (2009) measured intensity discrimination for a 65-dB SPL, 500-Hz pure tone. Thresholds in this task improved between 5 and 9 years of age, with adult-like performance in the oldest child listeners. Development effects in intensity discrimination spanning this age range have also been documented for stimuli consisting of a single pure tone (Maxon and Hochberg 1982), a three-tone complex (Willihnganz et al. 1997), and a narrow band of noise (Buss et al. 2006).

As in the frequency discrimination literature, there is some discrepancy regarding when in development intensity discrimination becomes adult-like. On the one hand, Berg and Boswell (2000) conclude that intensity discrimination is mature by 3 years of age provided the standard is sufficiently intense, whereas other studies indicate a prolonged development well into school age (e.g., Buss et al. 2009). A possible reason for this discrepancy is the role of memory and the use of gated versus continuous stimuli to estimate intensity discrimination. Detection of an increase in the level of a continuous tone (as in Berg and Boswell 2000) is associated with earlier maturation than detection of an increase in signal level of discrete stimulus presentations (as in Buss et al. 2009). However, this interpretation is undermined by the results of Fior and Bolzonello (1987). That study measured the ability to detect low-rate amplitude modulation in a continuous pure tone, a task sometimes referred to as Lüscher's test (Lüscher 1951), in a group of children who ranged from 3 to 10 years of age. There was a strong correlation between intensity discrimination threshold and listener age. This association was so striking in their data that Fior and Bolzonello (1987) proposed using this measure as an indicator of physiological maturation of hearing abilities.

4.3 Loudness

The perceptual correlate of intensity is loudness. Because loudness is subjective, it is difficult to estimate with confidence in infants and children. However, the data that are available indicate that adult-like loudness percepts develop early. Leibold and Werner (2002) demonstrated that reaction time in a detection task drops systematically with increasing intensity in a broadly similar way for 6- to 9-month-old infants and adults. While there was some indication of a steeper decline in reaction time in infants, it is unclear whether this trend reflects differences in loudness, as opposed to development of other processes underlying reaction time or increased variability in infant data. Collins and Gescheider (1989) asked children 4–7 years old to provide numerical estimates of the loudness of a 1,000-Hz tone and length of a line. Results of these two magnitude estimation tasks were consistent with results of a cross-modal task, in which line length was compared to tone loudness. Results in all conditions were stable over time for both groups, indicating that children are able to make consistent subjective estimates of loudness. Of particular interest here, results support the conclusion that children and adults have similar impressions of loudness. These results are broadly consistent with those of Serpanos and Gravel (2000), who tested 4- to 12-year-olds using a cross-modal task in which line length was adjusted to match the loudness of a tone or the level of a tone was adjusted to match the length of the line. The slope characterizing growth of loudness in that task was similar for children and adults, though there was a trend for children to judge sounds as being louder than adults. Macpherson et al. (1991) estimated the level associated with discomfort in normal hearing 3- to 5-year-olds. While some of the younger listeners were unable to perform the task, the data on the other listeners indicated a comparable or slightly elevated discomfort level relative to that in published studies with adults.

5 Simultaneous Masking and Frequency Resolution

One basic property of masking in adult listeners is that an increase in the level of a masking noise results in an analogous increase in the signal threshold. If the masking noise contains energy at the signal frequency and the masker and signal are presented simultaneously, then the signal threshold changes by 1 dB for every 1-dB change in masker level (Hawkins and Stevens 1950). Another important basic property of masking in adults is frequency selectivity, or how well the auditory system filters out masking energy that is separated in frequency from the frequency of the signal. A variety of measurement techniques, including behavioral masking (Fletcher 1940; Patterson 1976), auditory evoked potentials (Don and Eggermont 1978), and otoacoustic emissions (Kemp 1978), indicate that the adult human ear is highly frequency selective, probably due in large part to passive and active cochlear processes related to basilar membrane and outer hair cell function. Masked thresholds for a pure-tone signal and a masker with a relatively flat temporal envelope, such as wide bands of Gaussian noise, are thought to be limited by energetic masking; this masking is determined by peripheral frequency selectivity and efficiency, defined as the criterion signal-to-noise ratio (SNR) in the internal representation of the signal necessary for detection.

5.1 *Masked Threshold*

Several studies have shown a one-to-one relationship between detection threshold and the level of a simultaneous, on-frequency masker in infants and children. A study by Bull et al. (1981) showed that detection thresholds of 6- to 24-month-olds were approximately 15–25 dB higher than those of adults, but changing masker level resulted in a parallel effect in the two groups. The results of Schneider et al. (1990) and Bargones et al. (1995) were also consistent with infants having approximately 20-dB higher masked thresholds than adults. Results from Nozza and Wilson (1984) also indicated higher thresholds in infants (6–12 months old) than in adults, but only by approximately 6–8 dB. Whereas development of masked threshold may be related to frequency resolution, several authors have argued that peripheral frequency resolution is insufficient to account for threshold elevation in both infants (Spetner and Olsho 1990) and children (Schneider et al. 1990). The frequency effects observed for detection in quiet are sometimes also observed in masking noise: thresholds approach adult values earlier in development for high than low frequencies in some (Berg and Boswell 1999; He et al. 2010) but not all datasets (Schneider et al. 1989). When present, the frequency effect for masked detection appears to be more pronounced for brief than long duration tones (e.g., 10 vs. 100 ms) in both infants (Berg and Boswell 1999) and children (He et al. 2010).

5.2 *Frequency Resolution*

Otoacoustic emission data are consistent with grossly mature functioning of the active cochlear process in infants (Abdala 2001). It is therefore not surprising that some behavioral measures in infants are readily consistent with adult-like frequency selectivity. Olsho (1985) investigated frequency selectivity in infants and adults using a psychophysical tuning curve method wherein a fixed-frequency, low-level pure-tone signal was masked by an off-frequency pure tone of adjustable level. Good frequency selectivity is indicated when the level of the off-frequency maskers must be raised to relatively high values to mask the signal. Results showed comparable frequency selectivity in 5- to 8-month-olds and adults. A follow-up study using a nonsimultaneous pulsation threshold technique to estimate frequency selectivity confirmed that 6-month-olds had adult-like frequency selectivity, but found evidence of immature selectivity in 3-month-olds when tested at high frequencies (Spetner and Olsho 1990). A study by Schneider et al. (1990) that measured frequency selectivity using Fletcher's (1940) band-limited masking noise method also concluded that frequency selectivity was similar in adults and infants. In contrast, a study by Irwin et al. (1986) found results that were consistent with poorer frequency selectivity in 6-year-olds than in Fletcher's (1940) 10-year-olds and adults. In that study, a notched noise masking method (Patterson 1976) was used in which the level of a pure-tone signal was adjusted to obtain masked threshold in a fixed spectrum level masking noise as a function of the width of a spectral notch centered on the signal frequency. Using a similar notched noise method, Allen et al. (1989) found that frequency selectivity was reduced in 4-year-olds, but became adult-like by age 6 years. Both the Irwin et al. and Allen et al. studies showed higher overall masked thresholds in the children, a finding that was interpreted in terms of reduced listening efficiency. In agreement with Allen et al. (1989), Hall and Grose (1991) also found results consistent with reduced frequency selectivity for 4-year-olds using a notched noise method where signal threshold was estimated in a notched noise having a fixed spectrum level. However, they noted that the apparent poor frequency selectivity of the young listeners might arise due to effects related to poor listening efficiency and a relatively shallow growth of signal-related excitation in a notched noise masker. Using additional notched noise conditions designed to reduce factors related to the growth of excitation, Hall and Grose found similar frequency selectivity results between adult and 4-year-olds. This finding was consistent with the original conclusion of Olsho (1985) that the peripheral frequency selectivity is developed by infancy, and suggested that the poor performance of young listeners in fixed-level-masker notched noise maskers may be related to reduced processing efficiency.

Some masking effects wherein a signal of a given frequency must be detected in the presence of a masker at another frequency have been characterized in terms of distraction. For example, Werner and Bargones (1991) showed that a high-pass noise from 4,000 to 10,000 Hz had a significant masking effect on the detection of a 1,000-Hz tone in infants, even though there was likely to be negligible energetic masking for this stimulus and no masking in adult data. Supporting an interpretation

that the effect obtained in infants was not related to direct, energetic masking, the magnitude of the effect was the same whether the noise level was 40 dB SPL or 50 dB SPL. Results from Leibold and Neff (2007) suggested that school-age children can also show increased thresholds that are not likely to result from energetic masking for signals presented in the context of spectrally remote, fixed-frequency maskers. It is possible that this effect is related to the distraction masking reported in infants. Although such effects provide developmental information about how well the auditory system can filter out energy that is separated in frequency from the signal, they are likely to involve higher-level auditory processes than those that account for most basic frequency selective effects.

6 Nonsimultaneous Masking

Nonsimultaneous masking refers to an elevation in signal threshold brought about by a masker that precedes the signal in time (forward masking) or follows the signal in time (backward masking). In both cases, there is no physical overlap between the acoustic waveforms of the signal and masker. Werner (1999) assessed forward masking in 3- and 6-month-olds as a means of gauging maturation of adaptation. Using a 1,000-Hz probe tone masked by a broadband noise, she observed that for all intervals between the masker offset and the signal onset (Δt) infant thresholds were elevated relative to those of adults. However, the convergence on the unmasked signal threshold as Δt increased was relatively adult-like in the 6-month-olds but was more prolonged in the 3-month-olds. This suggests that infants younger than about 6 months of age are more susceptible to forward masking. Buss et al. (1999) found that forward masking performance continues to improve between 5 and 11 years of age. In that study, Δt was fixed at 0 ms and both narrow and broadband maskers were used.

The Buss et al. (1999) study also included a measure of backward masking for a Δt of 0 ms. Although performance in this condition was more variable than in other masking conditions tested, the effect of age was similar. This study also showed that thresholds in the backward masking condition were particularly susceptible to training effects. Hartley et al. (2000) also found that backward masking thresholds in 6-year-olds improved over successive measurement sessions. However, their analysis showed that this improvement was due to increasing listener age over the successive sessions and not to training, *per se*. Across the age span 6–10 years, they found a pronounced developmental effect of backward masking. This effect was confirmed in a later study by Ari-Even Roth et al. (2002) that included children up to age 11 years. Performance in this oldest group converged on that of adults.

The underlying reason for elevated backward masking thresholds in young children was addressed in a study by Hill et al. (2004). They concluded that the higher thresholds were not due to differences in temporal resolution—estimates of the temporal window were similar across listeners—but to poorer processing efficiency. Hill et al. (2004) sought to distinguish between processing efficiency and temporal

resolution by modeling the masking data in terms of temporal excitation patterns. These patterns are generated by the output of a model that includes a tapered temporal window, shaped by double-exponential functions (cf. Oxenham and Moore 1994). Based on parametric manipulation of the window shape, Hill et al. (2004) argued that the data were best accounted for by an age-invariant window, but that children required a higher relative signal level at the output of the window, akin to poorer processing efficiency.

It should be noted that the focus on backward masking in many of these studies was motivated by the report of Wright et al. (1997), who showed an apparently exacerbated effect of backward masking in children with specific language impairment. The topic of temporal processing characteristics, including backward masking performance, in special populations is beyond the scope of this chapter; the interested reader is referred to Bishop et al. (1999) as a starting point for further coverage.

7 Temporal Processing

Whereas Hill et al. (2004) held the temporal window constant and modeled age effects in terms of efficiency, other studies have pursued the idea that maturation of temporal processing may be associated with changes in the duration of the temporal window.

7.1 Temporal Integration

Several reports in the literature indicate that the level difference at threshold between a long-duration signal and a short-duration signal is more pronounced in infants than in adults (Berg 1991, 1993; Berg and Boswell 1995). This suggests that temporal integration is not adult-like in infancy. However, the dependence of this measure of temporal integration on signal frequency, level, and bandwidth in infants is complicated, and a simple picture does not emerge. Adding to the complexity is that if temporal integration is measured as the difference in level at threshold between a single tone-burst and a sequence of multiple tone-bursts (rather than between a single short stimulus and a single long stimulus) then no infant–adult differences are apparent (Berg and Boswell 1995).

Early data on integration in children provided only weak evidence of developmental effects. In a study of 4- to 12-year-old children, Maxon and Hochberg (1982) found no effect of age on the relative change in threshold as the duration of a signal varied between 25 ms and 800 ms, but they argued that the oldest children tested showed more integration than published adult data. In a clinically oriented study on brief-tone audiometry, Barry and Larson (1974) found relatively constant magnitudes of temporal integration across frequency in a group of children ages 6 to 14 years.

Although a later brief-tone audiometry study by Olsen and Buckles (1979) did note some differences among age subgroups within the overall range 6–24 years, these effects were nonmonotonic with respect to age and were interpreted in terms of other factors. More recently He et al. (2010) demonstrated a frequency effect for temporal integration in young children. In that study, 5- to 7-year-olds exhibited more temporal integration for a 1,625-Hz pure tone than older children or adults. In contrast, temporal integration was relatively uniform across 5- to 10-year-olds and adults for a 6,500-Hz pure tone. These interrelated effects of signal frequency and duration in masked detection are consistent with those demonstrated in infants (Berg and Boswell 1999).

7.2 *Temporal Resolution and Envelope Processing*

In contrast to temporal integration, temporal resolution refers to the acuity with which acoustic transitions in time can be followed. One of the most common gauges of temporal resolution is gap detection—the measurement of the briefest interruption in an ongoing sound that can be detected. A gap can be thought of as a single epoch of reduced amplitude in an otherwise steady envelope. Gap thresholds depend on various stimulus parameters such as bandwidth and level, spectral content of the markers bounding the gap, relative temporal position of the gap in the stimulus, etc., as well as the manner in which gap duration is defined. Even accounting for such parametric differences across studies, consensus on age effects is lacking. Irwin et al. (1985) found that gap thresholds are not adult-like until 10–12 years of age, as evident in data for both wide-band and more frequency-specific octave-band noises. In contrast, Wightman et al. (1989) found that children as young as 6 years of age exhibit adult-like gap detection for half-octave noise bands, with younger children (3–5 years) showing elevated thresholds. Complicating the picture further, maturation of gap detection appears to depend on particular parameters. For example, 8- to 10-year-old children exhibit adult-like gap thresholds in wide-band noise if the gaps are temporally positioned late relative to stimulus onset (50 ms), but their thresholds are poorer than adults' if the gaps occur soon after stimulus onset (5 ms; Diedler et al. 2007). Another way of assessing temporal resolution is to measure the temporal gap between two sounds required for the listener to hear two separate sounds rather than one continuous sound, a measure referred to as the point of auditory fusion threshold. Davis and McCroskey (1980) measured fusion thresholds in 3- to 12-year-old children and reported marked maturation between 3 and 8 years of age.

Whereas the age of convergence on adult-like performance for gap detection is in some dispute, there is little question that infants have markedly elevated thresholds. In infants, gap thresholds can be an order of magnitude higher than those of young adults (Werner et al. 1992). The elevation in gap thresholds is not likely to be due to a failure to encode the acoustic interruption because electrophysiological measures of gap detection using the auditory brain stem response (ABR) are similar between infants and adults (Werner et al. 2001). Moreover, the degree of threshold

elevation depends on the stimulus type. Trehub et al. (1995) employed brief 500-Hz tone-bursts where the overall stimulus shape was derived from overlapping Gaussian envelopes. They tested infants, 5-year-olds, and adults and found that gap thresholds improved across each age group; however, the actual threshold values were markedly lower than those found in other studies of similar age groups. Trehub et al. (1995) attributed this, in part, to a minimization of adaptation effects due to the use of brief stimuli.

Most of the developmental studies of gap detection have measured the ability to detect a gap in a stimulus with consistent spectral characteristics before and after that gap, a paradigm sometimes referred to as within-channel gap detection. In contrast, Smith et al. (2006) used an across-frequency-channel approach, where the leading gap marker was a tone-burst centered at 1,000 Hz and the trailing marker was a tone-burst centered at 4,000 Hz. They found that 6- to 7-month-old infants were significantly poorer than adults at detecting a gap between these two markers. Although this study did not test within-channel gap detection in the same infants, their stimuli were constructed in a similar manner to that of Trehub et al. (1995) described above. Smith et al. (2006) argued that across-channel gap detection appears to be less mature in 6-month-olds than within-channel gap detection.

A second psychophysical measure of temporal resolution based on sensitivity to dynamic envelope processing is the temporal modulation transfer function (TMTF). This paradigm gauges sensitivity to amplitude modulation as a function of modulation rate (Viemeister 1979). Available TMTF data indicate that 4- to 5-year-olds are less sensitive to modulation than 9- to 10-year-olds, and that children in the age range of 4–7 years of age are less sensitive than adults (Hall and Grose 1994). Adult-like TMTFs in older children (about 11–12 years) has been reported in a study by Lorenzi et al. (2000). Despite the poorer sensitivity observed by Hall and Grose (1994) in younger children, the shapes of the TMTFs were similar across all age groups. This result was interpreted as showing that the time constant defining temporal resolution of envelope encoding is independent of age, but that younger children are less efficient at processing the encoded information.

Whereas the TMTF measures the minimum detectable modulation depth as a function of modulation frequency, the masking period pattern is an envelope-processing task that makes use of suprathreshold levels of modulation depth. In this paradigm the detectability of a tonal signal is measured as a function of its temporal position relative to the modulation period of the masker. In an abbreviated form, the modified masker period pattern involves a simple comparison of masked threshold of a relatively long-duration signal presented in a modulated masker relative to an unmodulated masker (Zwicker and Schorn 1982). The benefit to tone detection of modulating the masker is an indication of the fidelity with which the masker envelope is being resolved. Grose et al. (1993) used the modified masking period pattern paradigm to estimate temporal resolution in 4- to 10-year-old children. They found that children exhibited adult-like temporal resolution by about 6 years of age for a signal frequency of 2,000 Hz. That is, 6-year-olds and adults benefited from masker modulation to a comparable degree at this frequency. In contrast, the magnitude of modulation benefit was still below adult levels at 10 years of age at a lower frequency region of 500 Hz.

7.3 *Temporal Fine Structure*

In contrast to envelope processing, temporal fine structure processing refers to the acuity with which cycle-by-cycle oscillations in sound pressure are followed in the audio frequency range.

Berg and Boswell (1995) proposed that immature ability to use temporal fine structure cues could be responsible for some of the age effects in their temporal integration data with 7-month-olds and adults. Whereas infant functions showing temporal integration of multipulse stimuli were adult-like by 7 months, age effects in temporal integration for a single pure-tone signal were observed for low (500 Hz) but not high (4,000 Hz) frequencies. Integration was greater in infants than in adults at 500 Hz, a result that Berg and Boswell hypothesize to be due to immature temporal fine structure coding. The idea that detection of brief, low-frequency stimuli relies on fine-structure cues is consistent with the finding that listeners with auditory nerve dys-synchrony are particularly poor at detecting very brief signals (Zeng et al. 2005).

Relatively few data speak to development of the ability to make use of temporal fine structure cues in monaural hearing. One indicator of the ability to make use of temporal fine structure cues is sensitivity to low-rate frequency modulation (FM; Moore and Sek 1996). Dawes and Bishop (2008) measured FM detection in school-age children and adults for a 500-Hz carrier frequency and both 2- and 40-Hz modulation rates. They found that children performed particularly poorly for the 2-Hz rate. This result is consistent with an inability of children to benefit fully from temporal fine-structure cues. However, data were not collected at a higher carrier frequency, a necessary control condition to rule out difficulty in detecting slow FM irrespective of fine-structure cues. The results of Bertoncini et al. (2009), on the other hand, argued that the ability to encode and use fine structure cues is fully developed by early childhood. That study assessed the importance of temporal fine structure for speech recognition using a paradigm in which the speech signal is digitally processed to remove the envelope, leaving only the frequency-modulating fine structure. Children as young as 5 years of age were able to discriminate this “temporal fine structure speech,” with performance comparable to that of adults.

One aspect of audition that relies on this fine-grained temporal processing in binaural hearing is the registration of interaural time differences (ITDs). Although the development of binaural processing is dealt with extensively by Litovsky (Chap. 6), it is touched on here because of its relevance to fine structure coding. The importance of binaural difference cues, such as ITDs, to spatial hearing is exemplified in the precedence effect wherein the spatial origin of a sound is driven primarily by the ITD associated with the leading wavefront (Blauert 1971). If the delay between the leading wavefront and any subsequent reflected wave (with its unique ITD) is sufficiently short, the echo is not perceived as a separate sound; only once the delay exceeds some critical value is the echo heard as a separate sound. The critical delay value depends on the type of sound being perceived. Morrongiello et al. (1984) measured the precedence effect in 6 month-old infants, 5-year-old children, and adults. Specifically, they measured the delay between the leading sound and the

echo necessary for the echo to be perceived as a separate sound. For click stimuli, children and adults performed similarly, whereas infants needed approximately twice the delay to respond to the echo. This result suggests a developmental progression either in the encoding of temporal fine structure or in the ability to make optimal use of fine structure cues.

7.4 *Duration Discrimination*

Acuity to variations in the temporal properties of sound can also be characterized by measuring the discrimination of durational changes. Morrongiello and Trehub (1987) measured duration discrimination in 6-month-olds, 5-year-olds, and adults using a rhythm change task, wherein the rhythm of a 7-s train of 200-ms noise-bursts was varied by shortening the durations of either the noise-bursts or the 200-ms interburst intervals during the middle section of the train. They found that the infants required a larger reduction in duration for accurate rhythm discrimination than did the children and, in turn, the children required a larger duration reduction than the adults did. However, across all age groups, rhythm discrimination did not depend on whether the duration reductions were applied to the noise-bursts or to the silent intervals between them. A later study by Elfenbein et al. (1993a) examined the development of duration discrimination in childhood. They measured sensitivity to both increments and decrements in the duration of a 350-ms noise-burst and found that these difference limens systematically decreased between 4 and 10 years of age, with performance of the 10-year-olds converging on that of adults. Interestingly, the younger children showed greater acuity for decrements in duration than for increments, a finding that was not evident in the performance of 10 year-olds and adults. Another finding of note was that the intersubject variability within age groups also decreased with increasing age. One facet of this greater spread in performance in the younger age groups was that, even among the 4-year-olds, there were individual listeners who provided difference limens on a par with those of adults.

A study by Jensen and Neff (1993a) around the same time generated findings similar to those of Elfenbein et al. (1993a). Jensen and Neff (1993a), measuring only duration increment discrimination for a 440-Hz, 400-ms tone, found a systematic improvement in difference limens between 4 and 6 years of age, with a concomitant reduction in intersubject variation within age groups. However, in this study, the best performance among 4-year-olds remained elevated compared to that of adults. Jensen and Neff (1993a) also measured frequency discrimination and intensity discrimination for the same standard stimulus, as described previously. Comparing the developmental effects across these three domains of discrimination, it was concluded that duration discrimination exhibited the most protracted developmental time course, with intensity discrimination exhibiting the earliest maturation and frequency discrimination being intermediary. In summary, duration discrimination improves throughout most of the childhood years, with even young children performing significantly better than infants.

8 Across-Channel Processing

Psychoacoustical studies to assess sensitivity to changes in basic features of sound, such as temporal or spectral processing, typically use very simple stimuli, such as pure tones. Natural stimuli are often spectrotemporally complex, however. Studies of across-channel processing shed light on some of the additional processes at work in complex stimulus perception.

8.1 Spectral Shape Discrimination

One cue that contributes importantly to the identification of speech, music, and other environmental sounds is the contour of the frequency spectrum. The processing of spectral contour cues has been studied via spectral shape discrimination and profile analysis (Green 1988) using paradigms that involve multitonal or noise stimuli. In such paradigms, good performance depends upon the processing of the *relative* amplitudes of the frequency components making up the stimulus because the sounds are manipulated in such a way that the absolute level at any frequency is an unreliable cue. Kidd et al. (1989) showed that the signal-to-noise ratio at which adults detect a pure tone in a noise background is very similar whether the noise is roved in level from trial to trial or is fixed in level. Kidd et al. noted that this could be accounted for by assuming that the listener can detect the signal in terms a change in the level profile across auditory filters. A study by Allen et al. (1998) indicates that children are also able to make use of across-channel cues in tone-in-noise detection tasks. In that study, a 1,000-Hz tone was detected in a band of noise from 800 to 1,200 Hz. Allen et al. found that 4- to 5-year-olds obtained similar detection thresholds whether or not the masker level was roved. This is consistent with the idea that adults and children have similar abilities to use spectral profile information in tone-in-noise detection tasks.

Not all paradigms are necessarily consistent with adult-like performance on spectral shape perception, however. Allen and Wightman (1992) examined spectral shape discrimination in 4- to 9-year-olds using a paradigm in which listeners were asked to discriminate a multitonal complex with a flat spectrum from the same complex with a sinusoidal spectral ripple. Thus, the task was not the detection of an added tone, but a change in the spectral shape of a tonal complex. The depth of sinusoidal spectral ripple was adaptively varied to obtain the discrimination threshold. The average results showed an improvement in sensitivity to spectral ripple from 4 to 9 years of age, with the results of some of the older listeners continuing to remain poorer than adult values. A notable finding was that, although the average performance of the children was often relatively poor, several of the listeners, even in the youngest group, showed sensitivity on a par with adult observers. Allen and Wightman (1992) suggested that this result could be interpreted as indicating that the cues for the detection of spectral ripple are well represented in the peripheral auditory system, even at a young age, but that the efficiency of the analysis of such cues can develop over the preschool and early school-age years, with individual differences in the rate of maturation of this ability.

8.2 *Comodulation Masking Release*

Comodulation masking release (CMR) refers to a detection advantage wherein a signal masked by a narrow band of noise centered on the signal frequency is made more detectable by the addition of flanking noise bands that share the amplitude fluctuation pattern of the on-signal noise band (Hall et al. 1984). In some CMR experiments, researchers have found that although school-aged children had elevated thresholds in the on-signal band baseline condition, they achieved adult-like release from masking (in dB) when comodulated flanking noise was added (Velooso et al. 1990; Grose et al. 1993; Hall et al. 1997). However, a study by Zettler et al. (2008) indicated that CMR increased in magnitude between 7 and 10 years of age. It is possible that procedural differences among the studies account for the disparity in outcome. For example, the frequency spacing among bands was smaller in the Zettler et al. study ($0.1 \times$ signal frequency) than the Velooso et al. study ($0.4 \times$ signal frequency) or the Hall et al. study ($0.25 \times$ signal frequency). Narrower frequency separations among noise bands appear to be associated with the introduction of within-channel cues (Schooneveldt and Moore 1987), and it is possible that the ability to process such cues changes with age.

The magnitude of CMR can be reduced by stimulus manipulations that diminish the extent to which the on-signal band is perceptually grouped with comodulated flanking bands (Hall and Grose 1990; Grose and Hall 1993). For example, CMR can be reduced by the inclusion of noise bands that have a different modulation pattern than the on-signal band and its comodulated flanking bands, particularly when the number of added bands is small (Hall and Grose 1990). Results from Hall et al. (1997) indicate that this effect is similar in adults and school-age children, a finding that is consistent with an interpretation that sound organization based upon patterns of across-frequency modulation may be similar in adults and children. The magnitude of CMR can also be reduced when the on-signal band and flanking bands are gated asynchronously (Grose and Hall 1993). This result has been interpreted as indicating that the CMR depends upon the on-signal band and flanking bands being processed as a single auditory object. One effect of asynchrony between the on-signal and flanking bands is to reduce the perceptual fusion of the bands. The reduction in CMR resulting from asynchrony between the on-signal and flanking bands is larger in school-age children than in adults (Hall et al. 1997). Hall et al. noted that this effect could indicate that children have some difficulty in processing comodulation in the context of a dynamic sequence of auditory objects. A discussion of the development of auditory grouping is found in Leibold (Chap. 5).

8.3 *Comodulation Detection Differences*

Comodulation detection differences (CDD) refer to a phenomenon involving the detection of a target narrow band of noise in the presence of one or more narrow bands of masking noise that are separated in frequency from the signal band

(McFadden 1987). When the masking and target bands of noise are comodulated, the threshold for the target band is higher (poorer) than when the target band has a different modulation pattern from the masking bands of noise. Although this phenomenon may be influenced by peripheral factors, such as energetic masking or suppression (Borrill and Moore 2002; Moore and Borrill 2002), part of the effect appears to be related to more complex factors associated with perceptual organization (McFadden 1987; Hall et al. 2006). For example, McFadden (1987) proposed that whereas the signal blends in with the masker in the comodulated condition, it emerges as a separate auditory object in the non-comodulated condition. Hall et al. (2008) found that CDD was much smaller in children 5–10 years of age than in adults (12.3 dB in adults vs. 3.5 dB in children). This finding suggests that grouping by common modulation is a relatively weak phenomenon in children, at least for the relatively brief (550 ms) stimuli used in the experiment (see also Leibold, Chap. 5). The results also indicated that the mechanisms giving rise to CDD appear to have a relatively protracted course of development, as the correlation between CDD and child age was not statistically significant over the wide age range tested.

9 Factors Responsible for Developmental Effects Observed Behaviorally

The protracted time course of auditory development across a broad range of paradigms could reflect a large number of factors, including anatomical, physiological, and cognitive maturation. As discussed in more detail by Abdala and Keefe (Chap. 2) and Eggermont and Moore (Chap. 3), the anatomy and physiology of the auditory system are relatively well developed at birth. The ear canal and middle ear continue to grow throughout childhood and early adolescence (Okabe et al. 1988), but the cochlea and neural periphery are grossly mature by 6 months of age (Abdala and Sininger 1996; Abdala 2001). Less is known about development of auditory processing proximal to the brain stem, but maturation of the auditory cortex continues through 12 years of age (Moore and Linthicum 2007), mirroring development of auditory and nonauditory attention (Gomes et al. 2000). Many previous behavioral studies of hearing have attempted to tease apart these effects with varying degrees of success (Schneider et al. 1989; Nozza 1995; Moore et al. 2008). This endeavor is complicated by individual differences unrelated to age and the possibility that multiple factors could affect individual listeners of the same age.

9.1 *Motivation, Attention, and Memory*

One troublesome aspect of estimating auditory sensitivity in infants and children is the relative inability to control or quantify factors related to attention and motivation. Whereas adults can provide feedback about their ability to maintain focus on the task, children are less able to do so. Schneider et al. (1989) have argued that the

relative stability of thresholds over time and the small effects of changing the task reward structure in adults indicate that motivation and inattention play little if any role in the poorer performance of child listeners. Some studies have evaluated the variance of listener responses over time in an attempt to quantify ability to maintain attention. For example, Wightman et al. (1989) tested the hypothesis that frequent, transient lapses in attention could explain developmental effects in a gap detection task with 3- to 7-year-olds and adults. Whereas the threshold estimation tracks in that study were characterized by comparable variance in child and adult listeners, thresholds tested on different days varied much more in child than in adult listeners' data. The possible effects of transient attention were modeled as a fixed proportion of random guesses that were unrelated to the stimulus, sometimes referred to as an "all-or-none" model because the listener is assumed to be either fully attentive or paying no attention at all. This model of inattention is associated with threshold elevation, reduction in psychometric function slope, and an asymptote below 100% correct. Although this model accounted for some of the variability in the data, Wightman et al. (1989) argued that this factor could not fully account for the developmental effects observed. The all-or-none model of inattention is often criticized as unrealistic (Schneider and Trehub 1992; Viemeister and Schlauch 1992; Allen and Wightman 1994), though at present there is no accepted alternative model of inattention. Perhaps the most compelling criticism of the hypothesis that child listeners respond randomly on some proportion of trials is the finding of differential trajectories of development for different auditory phenomena measured using common procedures (Jensen and Neff 1993b; Dawes and Bishop 2008).

In addition to central factors related to task orientation, the ability to store auditory stimuli in short-term memory could affect the ability to perform psychoacoustical tasks. Elfenbein et al. (1993b) suggested that the ability to store, retrieve and process short-term auditory memories could limit children's duration discrimination performance. Similar types of limits may also affect the ability to compare stimuli presented over time, such as in a three-alternative forced choice. Some evidence for a developmental effect in the duration of auditory sensory memory was reported by Gomes et al. (1999). That study used mismatched negativity (MMN), an event-related potential considered to be insensitive to attention, to assess frequency discrimination of a 1,000-Hz pulse train from a 1,200-Hz pulse train. Adults and 6- to 12-year-olds produced similar MMN responses when the pulse trains were separated by 1 s, but age-related differences in the MMN were noted when the trains were separated by 8 s. Gomes et al. (1999) interpreted this result as showing that the duration of auditory sensory memory develops over childhood, with some evidence of differences in the rate of development across individuals.

On the topic of individual differences, data from infants and children are nearly always more variable than those of adults. Within-listener variability could reflect measurement error or transient difficulties maintaining attention (Allen et al. 1989), but there is some evidence that individual differences may be relatively stable over time within an individual (e.g., Allen and Wightman 1994). To the extent that these differences are stable over time, they could reflect reliable individual differences in maturation. Further, it is sometimes argued that these differences are more consistent with differences in central processing than peripherally auditory processing.

9.2 *Detection in Quiet and Listening Behavior*

As discussed in the above, detection in quiet becomes adult-like earlier in development for high- than for low-frequency tones. This result is sometimes cited as evidence of development unrelated to attention or motivation, as central factors would be expected to apply equally to different frequency regions. Infants' pure-tone detection thresholds are more nearly adult-like at high than at low frequencies by 6 months of age (Olsho et al. 1988; Berg and Boswell 1999). This frequency-specific development can also be seen in thresholds of young children (Trehub et al. 1988). The frequency effect in tone detection is larger when tones are presented in quiet, but it is also evident in masking noise (Bargones et al. 1995; Berg and Boswell 1999). Whereas middle-ear maturation may affect detection in quiet, it would not be expected to play a role in masked detection for a masker that is clearly suprathreshold. The finding of frequency-specific developmental effects in masked detection is often cited as evidence that frequency-specific maturation is not due solely to peripheral effects, such as middle-ear transmission or self-generated acoustic noise.

Several authors have suggested that self-generated noise could play a role in frequency-specific development, particularly for detection in quiet (e.g., Bargones et al. 1995). When adults are trying to hear a soft sound, they engage in orienting behaviors that have been shown to improve performance (Stekelenburg and van Boxtel 2001), such as quiet breathing and reduced muscle activity. Because internal noise is predominantly low-frequency (Watson et al. 1972), Stekelenburg and van Boxtel (2001) hypothesized that orienting behaviors could have a larger effect on sensitivity for low-frequency than on high-frequency stimuli. This expectation received some support from the results of an experiment in which adults were asked to voluntarily contract facial muscles while listening, an activity that is thought to reduce auditory sensitivity by increasing the impedance of the middle ear. As expected, this manipulation increased detection threshold for a 100-Hz tone but had no effect on detection of a 1,500-Hz tone. While adults naturally engage in orienting behavior, there is some indication that this skill can be honed with practice, resulting in improved low-frequency thresholds (Zwislocki et al. 1958; Loeb and Dickson 1961). It is therefore possible that the frequency effect observed in infants and children could be related to maturation in the ability to engage in effective listening strategies.

9.3 *Masked Detection and Internal Noise*

Models of auditory detection often include two types of internal noise: additive noise, which forms the detection "floor" in quiet, and multiplicative noise, which limits performance for suprathreshold stimuli (sometimes also described in terms of efficiency). Some detection data from infants clearly reflect these two sources of noise, with differential threshold elevation in quiet and in the presence of a masker, and a fixed SNR across masker levels (Nozza 1995). In contrast, other studies report that the SNR at threshold falls with increasing masker level (Berg and Boswell 1999), a data pattern indicating a gradual transition between sensitivity in quiet and

in the presence of a masker. The requirement of increased SNR at threshold is sometimes described as reduced efficiency (Hall and Grose 1991, 1994; Hill et al. 2004). The factors responsible for reduced efficiency are not well understood, but are often thought to reflect central rather than peripheral processes.

One possible source of immaturity in masked detection is an inability to combine cues. Adding a tonal signal to a noise masker results in a number of cues, including an increase in signal energy and a reduction in stimulus envelope modulation depth. Whereas adults can make use of one or more different cues for a tone in noise task (Richards and Nekrich 1993), children may not be as flexible in this regard. For example, Allen et al. (1998) measured detection thresholds for a pure tone signal centered in a bandpass noise masker in 4- to 5-year-olds and adults. Moving the signal to the spectral edge of the masker improved thresholds for adults but not for children, and changing the signal to a band of noise hurt the performance of children but not adults. These results were interpreted as showing that children are not as good as adults at switching between or combining level, temporal, and spectral cues for tone-in-noise detection.

Allen and her colleagues have found some evidence that the psychometric function for masked tone detection is shallower and thresholds poorer for 3- to 5-year-olds than adults (Allen and Wightman 1994; Allen et al. 1998), though shallower slopes are not always found (Schneider et al. 1989). A possible explanation for shallower slopes and elevated thresholds in masked detection is that children are less adept than adults at focusing attention in the optimal frequency region. The idea that information distributed across frequency is combined nonoptimally in early childhood receives some support from Allen and Nelles (1996). Stimuli were tonal sequences with frequencies randomly drawn from a Gaussian distribution, and the task was to select the interval associated with the higher mean frequency. Children younger than 7 years of age performed more poorly than adults, particularly for longer stimulus sequences. This result was interpreted as showing that performance of the youngest listeners is limited by an inability to optimally combine frequency information distributed over time. Similar effects may be evident in infant data. Whereas adults tend to weight stimulus information more when it is in the frequency region of the expected signal, infants do not. This failure to listen in a frequency-specific way has been shown in infants using remote masking (Werner and Bargones 1991) and expected versus unexpected frequency paradigms (Bargones and Werner 1994). Failure to listen in a frequency-specific way could contribute to the difference in developmental results obtained with spectrally narrow versus broadband stimuli: 7- to 9-month-olds show greater threshold elevation in quiet and in masking noise when the target signal is a tone than when it is a broadband noise (Werner and Boike 2001). Frequency-specific listening and selective auditory attention are discussed in more detail by Leibold (Chap. 5).

9.4 Quantifying Internal Noise in Intensity Discrimination

One strategy for characterizing sources of immaturity is the use of psychophysical methods that support estimates of internal noise. In simple models of signal processing,

the slope of the psychometric function is a reflection of the variance in the cue underlying the decision process: the greater the variance, the shallower the function. Variance in the decision process comes from two sources, external noise based on stimulus variability and internal noise, related to the processing of that stimulus. If poor performance in young listeners is due to internal noise of this sort, then psychometric functions of infants and children should be shallower than adults', to the extent that internal noise limits performance. The observation of shallower psychometric functions for masked detection in children (Allen and Wightman 1994; Allen et al. 1998), noted above, is an example of this approach. Similarly, psychometric functions for detection in quiet in infants have been shown to become steeper with age (Olsho et al. 1988), with greater evidence of immaturity for short- than long-duration signals in 6- to 9-month-olds (Bargones et al. 1995).

The recent work on intensity discrimination in children presents evidence for a role of internal noise in auditory development of school-age children. Buss et al. (2006) measured intensity discrimination with and without stimulus level jitter, a form of external noise. Jitter tended to elevate thresholds in all listeners, but this effect was smaller for children than for adults. Reduced sensitivity to external noise was modeled as the result of a factor of 2.5 more internal noise in children, a finding that predicted a shallower psychometric function in children. Buss et al. (2009) tested this prediction of shallower slope in child listeners and found evidence that performance conforms closely to a simple model of signal detection, incorporating 2.5 times more internal noise for child than adult listeners. One assumption of this method is that performance of each listener is based on a stable and consistent listening strategy. Reliability and stability of listener strategy were evaluated in Buss et al. (2009) by assessing the correlation between two interleaved adaptive tracks (Leek et al. 1991). The correlation across pairs of interleaved tracks was no greater in child than in adult data, a result consistent with comparable stability in listener strategy over time for the two age groups (see also Allen et al. 1989; Wightman et al. 1989). Further, the relatively high upper asymptotes in the psychometric functions fitted to child data are consistent with sustained attention.

The results of Buss et al. (2006, 2009) do not identify the factors responsible for maturation of performance in the intensity discrimination task, but they do rule out some possibilities, notably transient inattention. Quantifying the developmental effect relative to the variance of the cue underlying adult performance also has the advantage of facilitating comparison across paradigms, including those with different task demands and stimuli. Such comparisons could expand our understanding of the factors responsible for auditory development.

10 Summary

Adaptive methods originally developed for testing adults, along with age-appropriate visual reinforcement techniques, have been highly effective in obtaining basic psychoacoustical data from infants and children. Although methodologically less efficient

than adaptive testing, fixed-block methods are very useful when the developmental question under study requires information about the steepness of the underlying psychometric function.

Although there are inconsistencies among studies about the ages at which performance on basic psychoacoustical tasks becomes adult-like, many auditory phenomena continue to show continued improvement at least through infancy and the early school-age years. One noteworthy trend, seen for threshold in quiet and in some aspects of temporal processing, is earlier development at higher than lower spectral regions. A challenge for future work is to determine the sources of such effect, including possible influences of frequency-specific self-generated noise or developmental effects related to tonotopically specific mechanisms of neural processing.

It is possible that the encoding of many aspects of sound at the sensory and peripheral neural levels of processing is relatively mature within the first year, and that improvements in auditory performance seen with increasing age reflect refinements in the ability of more central stages of processing to interpret information passed on by the periphery. This point of view is consistent, for example, with a derived measure of frequency resolution or temporal window time constant being similar between adults and children, while, at the same, children are poorer than adults in terms of raw measures of masked threshold and modulation detection.

The reasons underlying relatively poor performance by young listeners on basic psychoacoustic tasks are still poorly understood. Attempts to account for reduced performance in terms of lapses of attention have generally been unsuccessful. One potentially fruitful approach may be in terms of internal noise. This approach is closely tied to the slope of the psychometric function and may provide insights about the variance associated with the internal representation of the feature of the stimulus underlying performance. It is possible that a better understanding of the extent to which development across a range of auditory behaviors may be related can be achieved by analyzing similarities and differences in the internal noises associated with the various tasks.

References

- Abdala, C. (2001). Maturation of the human cochlear amplifier: Distortion product otoacoustic emission suppression tuning curves recorded at low and high primary tone levels. *Journal of the Acoustical Society of America*, 110(3 Pt 1), 1465–1476.
- Abdala, C., & Folsom, R. C. (1995). The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. *Journal of the Acoustical Society of America*, 98(2 Pt 1), 921–930.
- Abdala, C., & Sininger, Y. S. (1996). The development of cochlear frequency resolution in the human auditory system. *Ear and Hearing*, 17(5), 374–385.
- Allen, P., & Nelles, J. (1996). Development of auditory information integration abilities. *Journal of the Acoustical Society of America*, 100(2 Pt 1), 1043–1051.
- Allen, P., & Wightman, F. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research*, 35(1), 222–233.
- Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research*, 37(1), 205–215.

- Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research*, 32(2), 317–322.
- Allen, P., Jones, R., & Slaney, P. (1998). The role of level, spectral, and temporal cues in children's detection of masked signals. *Journal of the Acoustical Society of America*, 104(5), 2997–3005.
- Ari-Even Roth, D. A., Kishon-Rabin, L., & Hildesheimer, M. (2002). Auditory backward masking in normal hearing children. *Journal of Basic and Clinical Physiology and Pharmacology*, 13(2), 105–115.
- Aslin, R. N. (1989). Discrimination of frequency transitions by human infants. *Journal of the Acoustical Society of America*, 86(2), 582–590.
- Bargones, J. Y., & Burns, E. M. (1988). Suppression tuning curves for spontaneous otoacoustic emissions in infants and adults. *Journal of the Acoustical Society of America*, 83(5), 1809–1816.
- Bargones, J. Y., & Werner, L. A. (1994). Adults listen selectively; infants do not. *Psychological Science*, 5(3), 170–174.
- Bargones, J. Y., Werner, L. A., & Marean, G. C. (1995). Infant psychometric functions for detection: Mechanisms of immature sensitivity. *Journal of the Acoustical Society of America*, 98(1), 99–111.
- Barry, S. J., & Larson, V. D. (1974). Brief-tone audiometry with normal and deaf school-age children. *Journal of Speech and Hearing Disorders*, 39(457–464).
- Berg, K. M. (1991). Auditory temporal summation in infants and adults: Effects of stimulus bandwidth and masking noise. *Perception and Psychophysics*, 50(4), 314–320.
- Berg, K. M. (1993). A comparison of thresholds for 1/3-octave filtered clicks and noise bursts in infants and adults. *Perception and Psychophysics*, 54(3), 365–369.
- Berg, K. M., & Boswell, A. E. (1995). Temporal summation of 500-Hz tones and octave-band noise bursts in infants and adults. *Perception and Psychophysics*, 57(2), 183–190.
- Berg, K. M., & Boswell, A. E. (1998). Infants' detection of increments in low- and high-frequency noise. *Perception and Psychophysics*, 60(6), 1044–1051.
- Berg, K. M., & Boswell, A. E. (1999). Effect of masker level on infants' detection of tones in noise. *Perception and Psychophysics*, 61(1), 80–86.
- Berg, K. M., & Boswell, A. E. (2000). Noise increment detection in children 1 to 3 years of age. *Perception and Psychophysics*, 62(4), 868–873.
- Bertoncini, J., Serniclaes, W., & Lorenzi, C. (2009). Discrimination of speech sounds based upon temporal envelope versus fine structure cues in 5- to 7-year-old children. *Journal of Speech, Language, and Hearing Research*, 52(3), 682–695.
- Birnholtz, J. C., & Benacerraf, B. R. (1983). The development of human fetal hearing. *Science*, 222(4623), 516–518.
- Bishop, D. V., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research*, 42(6), 1295–1310.
- Blauert, J. (1971). Localization and the law of the first wavefront in the median plane. *Journal of the Acoustical Society of America*, 50(2B), 466–470.
- Borrill, S. J., & Moore, B. C. (2002). Evidence that comodulation detection differences depend on within-channel mechanisms. *Journal of the Acoustical Society of America*, 111(1 Pt 1), 309–319.
- Bull, D., Schneider, B. A., & Trehub, S. E. (1981). The masking of octave-band noise by broad-spectrum noise: A comparison of infant and adult thresholds. *Perception and Psychophysics*, 30(2), 101–106.
- Bull, D., Eilers, R. E., & Oller, D. K. (1984). Infants' discrimination of intensity variation in multisyllabic stimuli. *Journal of the Acoustical Society of America*, 76(1), 13–17.
- Buss, E., Hall, J. W., Grose, J. H., & Dev, M. B. (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, and Hearing Research*, 42(4), 844–849.
- Buss, E., Hall, J. W., Grose, J. H., & Dev, M. B. (2001). A comparison of threshold estimation methods in children 6–11 years of age. *Journal of the Acoustical Society of America*, 109(2), 727–731.
- Buss, E., Hall, J. W., & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *Journal of the Acoustical Society of America*, 120(5 Pt 1), 2777–2788.

- Buss, E., Hall, J. W., & Grose, J. H. (2009). Psychometric functions for pure tone intensity discrimination: Slope differences in school-aged children and adults. *Journal of the Acoustical Society of America*, 125(2), 1050–1058.
- Collins, A. A., & Gescheider, G. A. (1989). The measurement of loudness in individual children and adults by absolute magnitude estimation and cross-modality matching. *Journal of the Acoustical Society of America*, 85(5), 2012–2021.
- Davis, S. M., & McCroskey, R. L. (1980). Auditory fusion in children. *Child Development*, 51(1), 75–80.
- Dawes, P., & Bishop, D. V. (2008). Maturation of visual and auditory temporal processing in school-aged children. *Journal of Speech, Language, and Hearing Research*, 51(4), 1002–1015.
- Diedler, J., Pietz, J., Bast, T., & Rupp, A. (2007). Auditory temporal resolution in children assessed by magnetoencephalography. *NeuroReport*, 18(16), 1691–1695.
- Don, M., & Eggermont, J. J. (1978). Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *Journal of the Acoustical Society of America*, 63(4), 1084–1092.
- Elfenbein, J. L., Small, A. M., & Davis, J. M. (1993a). Developmental patterns of duration discrimination. *Journal of Speech and Hearing Research*, 36(842–849).
- Elfenbein, J. L., Small, A. M., & Davis, J. M. (1993b). Developmental patterns of duration discrimination. *Journal of Speech and Hearing Research*, 36(4), 842–849.
- Fior, R. (1972). Physiological maturation of auditory function between 3 and 13 years of age. *Audiology*, 11(5), 317–321.
- Fior, R., & Bolzonello, P. (1987). An investigation on the maturation of hearing abilities in children. *Ear and Hearing*, 8(6), 347–349.
- Fischer, B., & Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia*, 10(2), 105–118.
- Fletcher, H. (1940). Auditory patterns. *Review of Modern Physics*, 12(47–65).
- Folsom, R. C., & Wynne, M. K. (1987). Auditory brain stem responses from human adults and infants: Wave V tuning curves. *Journal of the Acoustical Society of America*, 81(2), 412–417.
- Gomes, H., Sussman, E., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan, H. G., Jr. (1999). Electrophysiological evidence of developmental changes in the duration of auditory sensory memory. *Developmental Psychology*, 35(1), 294–302.
- Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., & Cowan, N. (2000). The development of auditory attention in children. *Frontiers in Bioscience*, 5(D108–120).
- Gravel, J. S., & Traquina, D. N. (1992). Experience with the audiologic assessment of infants and toddlers. *International Journal of Pediatric Otorhinolaryngology*, 23(1), 59–71.
- Green, D. M. (1988). *Profile analysis: Auditory intensity discrimination*. New York: Oxford University Press.
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes-no task. *Journal of the Acoustical Society of America*, 93(4 Pt 1), 2096–2105.
- Grieco-Calub, T. M., Litovsky, R. Y., & Werner, L. A. (2008). Using the observer-based psychophysical procedure to assess localization acuity in toddlers who use bilateral cochlear implants. *Otology & Neurotology*, 29(2), 235–239.
- Grose, J. H., & Hall, J. W. (1993). Comodulation masking release: Is comodulation sufficient? *Journal of the Acoustical Society of America*, 93(5), 2896–2902.
- Grose, J. H., Hall, J. W., & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech and Hearing Research*, 36(2), 351–356.
- Hall, J. W., & Grose, J. H. (1990). Comodulation masking release and auditory grouping. *Journal of the Acoustical Society of America*, 88(1), 119–125.
- Hall, J. W., & Grose, J. H. (1991). Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation. *Journal of Speech and Hearing Research*, 34(3), 651–660.
- Hall, J. W., & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *Journal of the Acoustical Society of America*, 96(1), 150–154.

- Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, 76(1), 50–56.
- Hall, J. W., Grose, J. H., & Dev, M. B. (1997). Auditory development in complex tasks of comodulation masking release. *Journal of Speech, Language, and Hearing Research*, 40(4), 946–954.
- Hall, J. W., 3rd, Buss, E., & Grose, J. H. (2006). Comodulation detection differences for fixed-frequency and roved-frequency maskers. *Journal of the Acoustical Society of America*, 119(2), 1021–1028.
- Hall, J. W., 3rd, Buss, E., & Grose, J. H. (2008). Comodulation detection differences in children and adults. *Journal of the Acoustical Society of America*, 123(4), 2213–2219.
- Halliday, L. F., Taylor, J. L., Edmondson-Jones, A. M., & Moore, D. R. (2008). Frequency discrimination learning in children. *Journal of the Acoustical Society of America*, 123(6), 4393–4402.
- Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, and Hearing Research*, 43(6), 1402–1415.
- Hawkins, J. E., & Stevens, S. S. (1950). The masking of pure tones and of speech by white noise. *Journal of the Acoustical Society of America*, 22(1), 6–13.
- He, S., Buss, E., & Hall, J. W. (2010). Monaural temporal integration and temporally selective listening in children and adults. *Journal of the Acoustical Society of America*, 127(6), 3643–3653.
- Hicks, C. B., Tharpe, A. M., & Ashmead, D. H. (2000). Behavioral auditory assessment of young infants: Methodological limitations or natural lack of auditory responsiveness? *American Journal of Audiology*, 9(2), 124–130.
- Hill, P. R., Hartley, D. E., Glasberg, B. R., Moore, B. C., & Moore, D. R. (2004). Auditory processing efficiency and temporal resolution in children and adults. *Journal of Speech, Language, and Hearing Research*, 47(5), 1022–1029.
- Irwin, R. J., Ball, A. K., Kay, N., Stillman, J. A., & Rosser, J. (1985). The development of auditory temporal acuity in children. *Child Development*, 56(3), 614–620.
- Irwin, R. J., Stillman, J. A., & Schade, A. (1986). The width of the auditory filter in children. *Journal of Experimental Child Psychology*, 41(3), 429–442.
- Jensen, J., & Neff, D. (1993a). Development of basic auditory discrimination in preschool children. *Psychological Science*, 4(104–107).
- Jensen, J. K., & Neff, D. L. (1993b). Development of basic auditory discrimination in preschool children. *Psychological Science*, 4(104–107).
- Keller, T. A., & Cowan, N. (1994). Developmental increase in the duration of memory for tone pitch. *Developmental Psychology*, 30(6), 855–863.
- Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America*, 64(5), 1386–1391.
- Kidd, G., Jr., Mason, C. R., Brantley, M. A., & Owen, G. A. (1989). Roving-level tone-in-noise detection. *Journal of the Acoustical Society of America*, 86(4), 1310–1317.
- Leek, M. R., Hanna, T. E., & Marshall, L. (1991). An interleaved tracking procedure to monitor unstable psychometric functions. *Journal of the Acoustical Society of America*, 90(3), 1385–1397.
- Leek, M. R., Hanna, T. E., & Marshall, L. (1992). Estimation of psychometric functions from adaptive tracking procedures. *Perception and Psychophysics*, 51(3), 247–256.
- Leibold, L. J., & Werner, L. A. (2002). Relationship between intensity and reaction time in normal-hearing infants and adults. *Ear and Hearing*, 23(2), 92–97.
- Leibold, L. J., & Neff, D. L. (2007). Effects of masker-spectral variability and masker fringes in children and adults. *Journal of the Acoustical Society of America*, 121(6), 3666–3676.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(Supplement 2), 467–477.
- Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*, 117(5), 3091–3099.
- Loeb, M., & Dickson, C. (1961). Factors influencing the practice effect for auditory thresholds. *Journal of the Acoustical Society of America*, 33(7), 917–921.

- Lorenzi, C., Dumont, A., & Fullgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, 43(6), 1367–1379.
- Lüscher, E. (1951). The difference limen of intensity variations of pure tones and its diagnostic significance. *Journal of Laryngology and Otology*, 65(7), 486–510.
- Macpherson, B. J., Elfenbein, J. L., Schum, R. L., & Bentler, R. A. (1991). Thresholds of discomfort in young children. *Ear and Hearing*, 12(3), 184–190.
- Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing*, 3(6), 301–308.
- McFadden, D. (1987). Comodulation detection differences using noise-band signals. *Journal of the Acoustical Society of America*, 81(5), 1519–1527.
- Moore, B. C., & Sek, A. (1996). Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *Journal of the Acoustical Society of America*, 100(4 Pt 1), 2320–2331.
- Moore, B. C. J., & Borrill, S. J. (2002). Tests of a within-channel account of comodulation detection differences. *Journal of the Acoustical Society of America*, 112(5 Pt 1), 2099–2109.
- Moore, D. R., Ferguson, M. A., Halliday, L. F., & Riley, A. (2008). Frequency discrimination in children: Perception, learning and attention. *Hearing Research*, 238(1–2), 147–154.
- Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: A timeline of development. *International Journal of Audiology*, 46(9), 460–478.
- Moore, J. M., Thompson, G., & Thompson, M. (1975). Auditory localization of infants as a function of reinforcement conditions. *Journal of Speech and Hearing Disorders*, 40(1), 29–34.
- Moore, J. M., Wilson, W. R., & Thompson, G. (1977). Visual reinforcement of head-turn responses in infants under 12 months of age. *Journal of Speech and Hearing Disorders*, 42(3), 328–334.
- Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology*, 44(3), 413–426.
- Morrongiello, B. A., Kulig, J. W., & Clifton, R. K. (1984). Developmental changes in auditory temporal perception. *Child Development*, 55(2), 461–471.
- Nozza, R. J. (1995). Estimating the contribution of non-sensory factors to infant-adult differences in behavioral thresholds. *Hearing Research*, 91(1–2), 72–78.
- Nozza, R. J., & Wilson, W. R. (1984). Masked and unmasked pure-tone thresholds of infants and adults: Development of auditory frequency selectivity and sensitivity. *Journal of Speech and Hearing Research*, 27(4), 613–622.
- Okabe, K., Tanaka, S., Hamada, H., Tanetoshi, M., & Funai, H. (1988). Acoustic impedance measurement on normal ears of children. *Journal of the Acoustical Society of Japan*, 9(6), 287–294.
- Olsen, C. C., & Buckles, K. M. (1979). The effect of age in brief-tone audiometry. *Journal of Auditory Research*, 19(2), 117–122.
- Olsho, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior and Development*, 8(371–384).
- Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987a). Level and age effects in infant frequency discrimination. *Journal of the Acoustical Society of America*, 82(2), 454–464.
- Olsho, L. W., Koch, E. G., Halpin, C. F., & Carter, E. A. (1987b). An observer-based psychoacoustic procedure for use with young infants. *Developmental Psychology*, 23(5), 627–640.
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America*, 84(4), 1316–1324.
- Oxenham, A. J., & Moore, B. C. (1994). Modeling the additivity of nonsimultaneous masking. *Hearing Research*, 80(1), 105–118.
- Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America*, 59(3), 640–654.
- Primus, M. A. (1988). Infant thresholds with enhanced attention to the signal in visual reinforcement audiometry. *Journal of Speech and Hearing Research*, 31(3), 480–484.
- Primus, M. A. (1991). Repeated infant thresholds in operant and nonoperant audiometric procedures. *Ear and Hearing*, 12(2), 119–122.

- Richards, V. M., & Nekrich, R. D. (1993). The incorporation of level and level-invariant cues for the detection of a tone added to noise. *Journal of the Acoustical Society of America*, 94(5), 2560–2574.
- Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 3–46). Washington, DC: American Psychological Association.
- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1986). Auditory sensitivity in preschool children. *Journal of the Acoustical Society of America*, 79(2), 447–452.
- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America*, 86(5), 1733–1742.
- Schneider, B. A., Morrongiello, B. A., & Trehub, S. E. (1990). Size of critical band in infants, children, and adults. *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 642–652.
- Schooneveldt, G. P., & Moore, B. C. J. (1987). Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. *Journal of the Acoustical Society of America*, 82(6), 1944–1956.
- Serpanos, Y. C., & Gravel, J. S. (2000). Assessing growth of loudness in children by cross-modality matching. *Journal of the American Academy of Audiology*, 11(4), 190–202.
- Shahidullah, S., & Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Human Development*, 36(1), 13–26.
- Sinnott, J. M., & Aslin, R. N. (1985). Frequency and intensity discrimination in human infants and adults. *Journal of the Acoustical Society of America*, 78(6), 1986–1992.
- Smith, N. A., Trainor, L. J., & Shore, D. I. (2006). The development of temporal resolution: Between-channel gap detection in infants and adults. *Journal of Speech, Language, and Hearing Research*, 49(5), 1104–1113.
- Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, 61(3), 632–652.
- Stekelenburg, J. J., & van Boxtel, A. (2001). Inhibition of pericranial muscle activity, respiration, and heart rate enhances auditory sensitivity. *Psychophysiology*, 38(4), 629–641.
- Sutcliffe, P., & Bishop, D. (2005). Psychophysical design influences frequency discrimination performance in young children. *Journal of Experimental Child Psychology*, 91(3), 249–270.
- Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41(4A), 782–787.
- Tharpe, A. M., & Ashmead, D. H. (2001). A longitudinal investigation of infant auditory sensitivity. *American Journal of Audiology*, 10(2), 104–112.
- Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *Journal of Speech, Language, and Hearing Research*, 42(5), 1061–1068.
- Trehub, S. E., Schneider, B. A., & Bull, D. (1981). Effect of reinforcement on infants' performance in an auditory detection task. *Developmental Psychology*, 17(6), 872–877.
- Trehub, S. E., Schneider, B. A., Morrongiello, B. A., & Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *Journal of Experimental Child Psychology*, 46(2), 273–285.
- Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustical Society of America*, 98(5 Pt 1), 2532–2541.
- Veloso, K., Hall, J. W., 3rd, & Grose, J. H. (1990). Frequency selectivity and comodulation masking release in adults and in 6-year-old children. *Journal of Speech and Hearing Research*, 33(1), 96–102.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *Journal of the Acoustical Society of America*, 66(5), 1364–1380.
- Viemeister, N. F., & Schlauch, R. S. (1992). Issues in infant psychoacoustics. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 191–209). Washington, DC: American Psychological Association.

- Watson, C. S., Franks, J. R., & Hood, D. C. (1972). Detection of tones in the absence of external masking noise. I. Effects of signal intensity and signal frequency. *Journal of the Acoustical Society of America*, 52(2B), 633–643.
- Werner, L. A. (1999). Forward masking among infant and adult listeners. *Journal of the Acoustical Society of America*, 105(4), 2445–2453.
- Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception and Psychophysics*, 50(5), 405–412.
- Werner, L. A., & Boike, K. (2001). Infants' sensitivity to broadband noise. *Journal of the Acoustical Society of America*, 109(5 Pt 1), 2103–2111.
- Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development*, 63(2), 260–272.
- Werner, L. A., Folsom, R. C., Mancl, L. R., & Syapin, C. L. (2001). Human auditory brainstem response to temporal gaps in noise. *Journal of Speech, Language, and Hearing Research*, 44(4), 737–750.
- Werner, L. A., Parrish, H. K., & Holmer, N. M. (2009). Effects of temporal uncertainty and temporal expectancy on infants' auditory sensitivity. *Journal of the Acoustical Society of America*, 125(2), 1040–1049.
- Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development*, 60(3), 611–624.
- Willihnganz, M. S., Stellmack, M. A., Lutfi, R. A., & Wightman, F. L. (1997). Spectral weights in level discrimination by preschool children: Synthetic listening conditions. *Journal of the Acoustical Society of America*, 101(5 Pt 1), 2803–2810.
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, 387(6629), 176–178.
- Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology*, 93(6), 3050–3063.
- Zettler, C. M., Sevcik, R. A., Morris, R. D., & Clarkson, M. G. (2008). Comodulation masking release (CMR) in children and the influence of reading status. *Journal of Speech, Language, and Hearing Research*, 51(3), 772–784.
- Zwicker, E., & Schorn, K. (1982). Temporal resolution in hard-of-hearing patients. *Audiology*, 21(6), 474–492.
- Zwislocki, J., Maire, F., Feldman, A. S., & Rubin, H. (1958). On the effect of practice and motivation on the threshold of audibility. *Journal of the Acoustical Society of America*, 30(4), 254–262.

Chapter 5

Development of Auditory Scene Analysis and Auditory Attention

Lori J. Leibold

1 Introduction

This chapter summarizes the literature describing auditory scene analysis and auditory attention during infancy and childhood. In Chap. 4, Buss, Hall, and Grose reviewed more than 30 years of research investigating the accuracy with which the frequency, intensity, and temporal characteristics of incoming sounds are represented by the developing child. Building on these impressive studies, a basic assumption underlying the work described in the current chapter is that the peripheral auditory system provides the brain with a precise representation of most sounds by at least 6 months after term birth.

Despite the precocious development of the sensory representation of sound, it is clear that many aspects of hearing follow a prolonged time course of development. For example, the ability to extract and selectively attend to target acoustic information in the presence of competing background sounds such as speech remains a difficult challenge for many adolescents (e.g., Wightman and Kistler 2005). These prolonged immaturities in hearing are widely thought to reflect age-related changes in perceptual processing operating within the central auditory system. This processing includes the ability to group acoustic waveforms into different sounding objects (auditory scene analysis) as well as the ability to select the appropriate auditory object for further processing while discounting information not relevant to the task (selective attention).

L.J. Leibold (✉)

Department of Allied Health Sciences, The University of North Carolina
at Chapel Hill, School of Medicine, Chapel Hill, NC, USA
e-mail: leibold@med.unc.edu

The information covered by this chapter is organized into two main sections. The first section describes age-related changes in the ability to perform auditory scene analysis, including studies of both auditory stream segregation and informational masking. The second section deals with the development of auditory attention, with a particular focus on studies that have examined the developing child's ability to selectively attend to a target sound in the presence of competing maskers or attend to a particular feature within a complex sound. Although the chapter is organized according to these two distinct sections, it is important to point out that it is unlikely that these processes operate independently. In addition, it can be difficult to distinguish between these two processes during infancy and childhood.

2 Auditory Scene Analysis

Infants and children listen to and learn about important sounds such as speech in natural acoustic environments. These environments typically contain multiple sources of competing background sounds. Moreover, each source can produce acoustic waveforms consisting of many individual frequencies that change over time. Reaching the ears of the developing child is an overlapping mixture of waveforms produced by all of the active sources in their environment. For example, a young child at home may try to follow his father's voice as he reads him a story. At the same time, however, sounds are produced by other sources in the home such as a television and the boy's older sister who is playing the guitar. The challenge for this child is that, to hear and understand the story, he must be able to group the frequency components that were produced both sequentially and simultaneously from his father's voice as separate from those produced by the television and the guitar. The general process by which the brain performs this task of parsing acoustic waveforms and assigning them to the appropriate source has been referred to in the literature as auditory scene analysis (Bregman 1990).

Bregman (1990) proposed auditory scene analysis as a process whereby listeners segregate incoming acoustic waveforms into multiple sources of sound. Bregman's theory includes an early or "primitive" stage based on segregating incoming acoustic waveforms on the basis of general acoustic characteristics and regularities inherent in sounds produced by the same source. It is generally assumed that listening experience is not required to segregate sounds using these primitive mechanisms. Auditory scene analysis also consists of a later "schema-based" stage, relying on learned schemas that are dependent on a listener's previous experiences with sound. Note that most studies of adults and nearly all developmental studies described in this chapter focused on identifying and manipulating acoustic cues thought to be involved in the primitive grouping mechanisms. These mechanisms are appealing to study during infancy and childhood because the listener requires no previous listening experience with particular sounds to segregate them. In addition, these mechanisms could allow a developing auditory system to be refined over time.

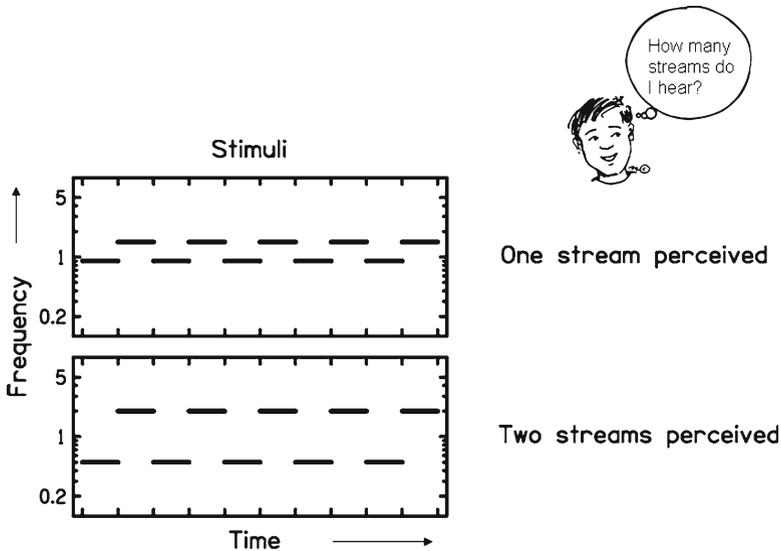


Fig. 5.1 A schematic illustration of auditory stream segregation. The child perceives a single coherent stream of tones in the *top panel* when the frequency separation between successive tones is small. Two distinct streams are perceived by the same child in the *bottom panel* when the frequency separation between successive tones is increased

2.1 Auditory Stream Segregation

Auditory streaming methods are commonly used to study auditory scene analysis. Auditory stream segregation refers to the ability to group incoming waveforms into separate auditory streams on the basis of acoustic cues that promote coherence across time. In a typical streaming paradigm, listeners are presented with a relatively simple sound sequence in which elements of the sequence alternate between two values within a particular acoustic dimension. For example, Fig. 5.1 shows a schematic representation of a temporal sequence of two pure tones differing only in frequency. In this example, the frequency separation between the alternating pure tones is manipulated, and listeners report whether they heard a single coherent stream or two distinct streams. If the frequency difference between the tones is small, listeners report hearing a single stream with a galloping rhythm. In contrast, listeners report hearing two concurrent streams of tones if the frequency difference between the tones is large (e.g., van Noorden 1975).

Research conducted during the past several decades has identified a variety of acoustic cues that promote the formation of auditory streams. This research has demonstrated that adults can separate streams of incoming sounds on the basis of spatial separation, frequency separation, spectral profile, talker sex, onset and offset times, temporal modulation, and harmonicity (reviewed by Bregman 1990; Darwin and Carlyon 1995; Yost 1997). For example, common temporal onsets and offsets

appear to be robust cues for performing auditory stream segregation. That is, frequency components that start and stop at the same time are more likely to have originated from the same sound source compared to frequency components that start or stop at different times (e.g., Bregman and Pinker 1978).

2.2 *Auditory Stream Segregation in Infants*

The ability to form auditory streams has been examined during infancy, but the data are limited due largely to methodological difficulties associated with designing a feasible infant test paradigm. For example, infants cannot directly indicate whether they perceive a sound sequence as a single auditory stream or as multiple streams. In addition, verbal instructions cannot be effectively used to direct infants' behavior. Thus, indirect discrimination paradigms have been developed to study auditory streaming in infants. These paradigms are based on research demonstrating that the formation of auditory streams can affect adults' ability to judge the order of tones presented rapidly in a sequence (e.g., Broadbent and Ladefoged 1959; Warren et al. 1969; Bregman and Campbell 1971). For example, Bregman and Campbell (1971) presented adults with a repeating sequence of three high-frequency and three low-frequency pure tones. The high- and low-frequency tones were interleaved. Listeners had difficulty judging relative tone order in the sequence across the high- and low-frequency tones, but could accurately report the relative order of either the high- or the low-frequency tones separately. The investigators concluded that these findings reflected the fact that the low and high tones were separated into different auditory streams. That is, listeners could not judge the order of the tones when the tones were perceived as being in distinct streams.

Following the general approach outlined in the preceding text for adults, Demany (1982) examined whether 2–4-month-old infants and adults were similar in how they organized auditory streams on the basis of frequency separation. A repeating sequence of four different pure tones was presented in either forward or reverse relative order. As with the previous adult studies described above, the assumption was that the reverse sequence would be discriminated from the forward sequence if adjacent tones were assigned to the same auditory stream, but not if they were assigned to different streams. Whereas the formation of streams was determined using same/different judgments for adults, infants were tested by measuring looking time in the context of a habituation/dishabituation paradigm. The forward sequence was repeated until the infant habituated and then the reverse sequence was presented. Infants' discrimination was inferred from the recovery of looking time to the reversed sequence. Infants and adults appeared to discriminate similar changes in the relative tone order of sequences, suggesting that infants organize auditory streams on the basis of frequency proximity in a manner similar to adults.

Methodological issues limit interpretation of the results reported by Demany (1982). Specifically, acoustic cues related to the frequency contour of the sequences were likely sufficient to allow listeners to discriminate the forward and reversed tone

sequences, regardless of whether or not the sequences were perceived as one or two streams. This problem was addressed in a subsequent study conducted by Fassbender (1993), who examined 2–5-month-olds' ability to discriminate sequences that adults organize on the basis of frequency range, amplitude, or timbre using a nonnutritive sucking habituation paradigm. Similar to Demany (1982), infants appeared to discriminate changes in the temporal order of the sequences similarly to adults.

Finally, McAdams and Bertoncini (1997) examined the discrimination abilities of 3–4-day-old infants on sequences that adults segregate on the basis of both spatial position and timbre. Following Demany (1982), repeating four-tone patterns were repeated in either a forward or a reverse sequence. A habituation/dishabituation paradigm of nonnutritive sucking was used to test the infants. Both infants and adults were unable to discriminate the forward from the reverse sequences, suggesting that alternating tones were perceived as separate streams. Note, however, that it could not be determined whether infants organized the sequences by spatial location, timbre, or both.

In summary, results across infant habituation/dishabituation studies are consistent with the hypothesis that the ability to perform sequential auditory stream segregation is present early during infancy. Moreover, infants appear to use some of the same acoustic cues used by adults to segregate auditory streams (Demany 1982; Fassbender 1993; McAdams and Bertoncini 1997). Nonetheless, caution is warranted when interpreting the results obtained from these initial studies. Although the process of auditory stream segregation appears functional early in life, it is not clear how accurately this process operates during infancy. The infant studies inferred auditory streaming abilities based on discrimination performance using relatively large acoustic differences. In addition, procedural modifications introduced to test infants limit direct comparisons with adult data.

2.3 Auditory Stream Segregation in Preschoolers and School-Age Children

For many psychoacoustic abilities, procedures similar to conventional adult paradigms can be used to test children by about 4–5 years of age. Thus, it is surprising that few studies have examined the auditory streaming abilities of preschoolers and school-age children. One exception is a study by Sussman et al. (2007), who compared the frequency separation required to form auditory streams across younger children (5–8 years), older children (9–11 years), and adults. Using a conventional auditory streaming paradigm, listeners were presented with a sequence of two alternating pure tones. One tone was fixed in frequency and the other tone was varied in frequency, resulting in five different frequency separations across testing conditions. Listeners were asked to report whether they heard one or two streams. Compared to adults and older children (9–11 years), younger children (5–8 years) required a significantly larger frequency separation between alternating tones before indicating they heard two distinct auditory streams. These results are consistent

with the idea that the process of auditory stream segregation continues to be refined into the school-age years.

Indirect information regarding auditory stream segregation during childhood can also be found in the literature describing the development of comodulation masking release (CMR). CMR refers to the reduction in the masked threshold of a signal when intensity fluctuations of remote-frequency or “flanking” maskers correlate with intensity fluctuations of a masker centered on the signal frequency (Hall et al. 1984). That is, the addition of remote-frequency noise improves listeners’ detection thresholds when listeners are able to benefit from the similarities in amplitude modulation across multiple auditory filters. Presumably, the flanking and on-signal maskers are perceptually grouped together as a single stream distinct from the signal.

Results from studies of CMR in children suggest that the basic mechanisms responsible for CMR are adult-like by 4–5 years of age (e.g., Grose et al. 1993; Veloso et al. 1990). For example, Veloso et al. (1990) compared the performance of 6-year-old children and adults in the detection of a 1,000-Hz signal in the presence of one on-signal and two flanking narrow bands of noise. Although masked thresholds were uniformly higher for children than for adults, a similar CMR was observed between the two age groups when the amplitude envelopes of the flanking and on-signal noise bands correlated compared to when the noise bands did not correlate. These findings suggest that children can benefit from auditory grouping cues related to the coherence of modulation patterns across frequency by kindergarten. Note, however, that an extended time course of development has been reported for some complex CMR conditions (e.g., Hall et al. 1997; Zettler et al. 2008). Hall et al. (1997) found that the CMR was often eliminated or made negative for 5–11-year-old children in conditions in which there was a temporal asynchrony between the on-signal and flanking bands of masker noise. The CMR for adults was reduced, but generally not eliminated, with the introduction of a temporal asynchrony.

2.4 Summary and Future Directions

Results from indirect studies of auditory stream segregation in infants indicate that the basic mechanisms are functional early in life and that even newborns use many of the same acoustic cues that adults use to segregate auditory streams (Demany 1982; Fassbender 1993; McAdams and Bertoncini 1997). Results from studies investigating auditory streaming and CMR in preschoolers and older children extend the infant work, but suggest that the time course of development for reaching the accuracy shown by adults for at least some auditory streaming tasks extends well into the school years.

To date, most investigations of auditory stream segregation involving infants and children examined performance using simple pure-tone stimuli that did not overlap in time. One area of interest for future studies is to characterize the developmental trajectory of auditory stream segregation using more complex and natural sounds that require the listener to segregate sound sources both sequentially and simultaneously.

A second area of considerable interest is to determine which acoustic cues are the most salient during development and whether infants and children rely on the presence of multiple or redundant acoustic cues to perform sound source segregation. Bregman (1993) has argued that multiple strategies and auditory grouping cues are available to adult listeners in most natural listening environments, but it is yet unclear whether infants and children rely on the same acoustic cues as adults or can integrate information across multiple acoustic cues to improve their ability to segregate sounds in noisy environments.

3 Informational Masking

An appealing approach for studying the development of auditory scene analysis is to examine whether acoustic cues believed to aid in sound source segregation provide a release from “informational masking” (e.g., Watson et al. 1975; Neff and Green 1987; Kidd et al. 1994). Whereas energetic masking is believed to reflect peripheral limitations of the auditory system (Fletcher 1940), informational masking generally refers to masking produced even though the peripheral auditory system provides the central auditory system with sufficient frequency, temporal, and intensity resolution to encode both the target and the masker. Although a precise operational definition of informational masking across all stimuli and measurement procedures remains elusive, a comprehensive description provided by Brungart (2005) states that “informational masking occurs when the signal and masker are both audible but the listener is unable to disentangle the elements of the target signal from a similar-sounding distractor” (p. 261).

3.1 Informational Masking Paradigms

A number of sophisticated paradigms have been established to examine informational masking in adults (e.g., Watson et al. 1975; Neff and Green 1987; Kidd et al. 1994), and several of these paradigms have been applied to studies involving infants and children. Perhaps the most widely used approach is the “simultaneous multitonal masking paradigm” developed by Neff and Green (1987). A schematic example is provided in the top panel of Fig. 5.2. In this example, the listener’s task is to detect a 1,000-Hz tone (red line) presented simultaneously with a multitonal masker comprised of four components (gray lines). As in the original study by Neff and Green, the frequency components were randomly varied on each presentation within the range of 300–3,000 Hz. To limit energetic masking, components did not within a 160-Hz frequency range centered on the 1,000-Hz signal. Neff and Green observed masking effects as large as 50 dB for some listeners when the random-frequency masker comprised ten components. Atypical of energetic masking, individual differences on the order of 40 dB were reported.

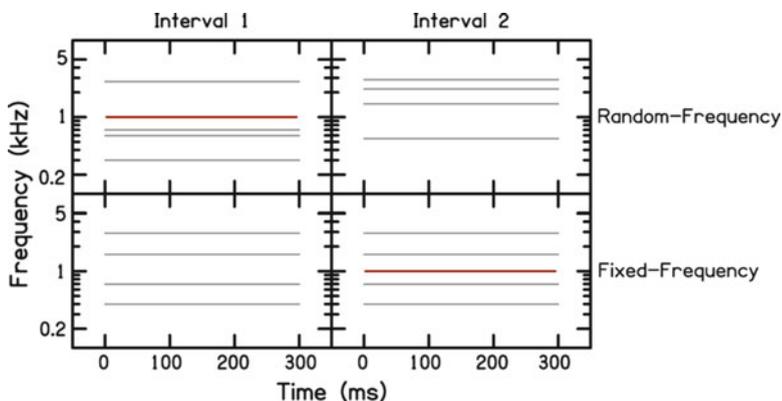


Fig. 5.2 A schematic representation of the simultaneous multitonal masking paradigm developed by Neff and Green (1987). This example shows two intervals of a 2IFC trial for both random- and fixed-frequency masker conditions in the form of spectrograms. In the *top panels*, the frequency components comprising the four-tone masker were selected randomly on each presentation (random-frequency condition). In the *bottom panels*, the frequency components comprising the four-tone masker were fixed in frequency across listening intervals (fixed-frequency condition). The masker components are shown by gray lines. The 1,000-Hz signals are shown by the red lines added to one of the two listening intervals in each panel

Later studies of informational masking in adults using the simultaneous multitonal masker paradigm confirmed that randomizing the spectral content of a multitonal masker on each presentation can produce substantial masking for many adults and large individual differences across listeners (e.g., Oh and Lutfi 1998; Wright and Saberi 1999; Richards et al. 2002; Durlach et al. 2005). Investigators have also reported a nonmonotonic relation between detection performance and the number of components comprising the random-frequency multitonal masker (e.g., Neff and Green 1987; Oh and Lutfi 1998). When components are added to a masker with a relatively small number of components (10–20), more masking is produced. In contrast, adding components to a masker with 20 or more components typically results in reduced masking. This finding is believed to reflect greater contributions of informational masking for maskers with relatively few components and increased contributions of energetic masking as the density of masker components approaches broadband noise.

Early adult studies of informational masking often examined effects of masker-spectral uncertainty, but researchers have also reported that thresholds in conditions with little or no masker-frequency uncertainty (e.g., masker samples fixed across intervals of each trial or across the entire block of trials) are often considerably higher than the absolute threshold for the signal in quiet (e.g., Neff and Callaghan 1988; Neff and Dethlefs 1995; Wright and Saberi 1999; Alexander and Lutfi 2004; Richards and Neff 2004; Durlach et al. 2005; Leibold et al. 2010). For example,

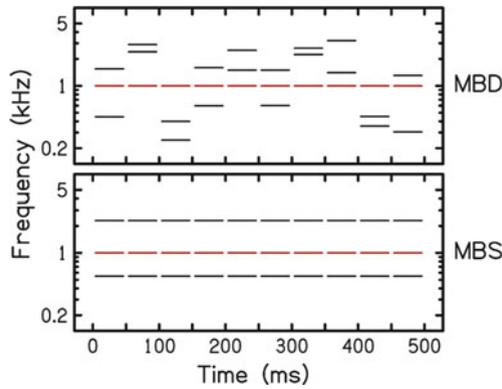


Fig. 5.3 A schematic representation of the multiple-bursts paradigm developed by Kidd et al. (1994). In this example, the task is to detect a sequence of 10 bursts of a 1,000-Hz signal (*red lines*) in the presence of a sequence of 10 bursts of a two-tone masker (*black lines*). In the *top panel*, masker components were chosen randomly for each burst in the sequence (*MBD condition*). In the *bottom panel*, masker components are fixed across bursts within a given listening interval (*MBS condition*)

Durlach et al. (2005) reported average thresholds for a 1,000-Hz signal ranging from 33 to 48 dB SPL across ten listeners in the presence of an eight-tone, fixed-frequency, simultaneous masker. These results are consistent with the idea that stimulus uncertainty is not required to produce informational masking (e.g., Durlach et al. 2003). A schematic example of a fixed-frequency multitonal masker sample is shown in the bottom panel of Fig. 5.2.

The “multiple-bursts paradigm,” developed by Kidd et al. (1994), has also been applied to the study of informational masking in school-age children (e.g., Hall et al. 2005; Leibold and Bonino 2009). An illustration of stimuli used in a typical multiple-bursts experiment is shown in Fig. 5.3. In this example, the listener’s task is to detect a sequence of repeating 1,000-Hz signal bursts (red lines) embedded in a sequence of multitonal masker bursts (black lines). The top panel of Fig. 5.3 shows a multiburst-different (MBD) condition in which masker components are chosen randomly for each burst. The bottom panel shows a multiburst-same (MBS) condition in which masker components are fixed across bursts within a given listening interval. When present, the pure-tone signal bursts are gated synchronously with the masker tones. Contrary to the simultaneous multitonal masking paradigm (Neff and Green 1987), detection performance is typically better in MBD compared to MBS maskers. That is, the introduction of masker uncertainty results in less masking compared to when the same masker sample is fixed across presentations. The interpretation of this result is that the spectro-temporal constancy of the signal bursts contrast with the random-frequency masker bursts that comprise the MBD masker to form a distinct and coherent auditory stream. In contrast, the MBS masker bursts and signal tones are often perceived as a single auditory object (e.g., Kidd et al. 1995; Huang and Richards 2006).

3.2 *Developmental Effects in Susceptibility to Informational Masking*

On average, infants and children are more susceptible to informational masking than adults (e.g., Allen and Wightman 1995; Oh et al. 2001; Wightman et al. 2003; Leibold and Werner 2006). Similar to the early adult work, initial developmental studies of informational masking examined effects of masker-frequency uncertainty on children's detection performance. For example, Allen and Wightman (1995) measured detection thresholds for a 1,000-Hz tone in the presence of broadband noise alone and in the presence of the same broadband noise plus a single, random-frequency pure tone. Listeners were 3–5-year-old children and adults. Although large individual variability was observed, children generally performed much worse than adults. Whereas the group average threshold for children was 81 dB SPL in the combined noise and random tonal masker condition, the group average threshold for adults in this condition was 66.2 dB SPL. Note also that the data of only 7 of 17 preschoolers tested were used to compute the group average child threshold. The remaining ten children did not have measureable thresholds less than 90 dB SPL.

Oh et al. (2001) reported additional data describing the informational masking produced by random-frequency multitonal maskers in preschoolers and adults. Using the simultaneous multitonal masking paradigm (Neff and Green 1987), listeners were asked to detect a 1,000-Hz pure tone presented simultaneously with either a broadband noise or a random-frequency, multitonal masker. The number of masker components was varied across experimental conditions, ranging from 2 to 906. Group average data are shown in Fig. 5.4. Consistent with previous studies of adults (e.g., Neff and Green 1987; Oh and Lutfi 1998), a nonmonotonic relation between amount of masking and number of masker components was observed for both children and adults. Although greater masking was observed for children compared to adults for all masker conditions, the developmental effect was most pronounced for random-frequency maskers with relatively few components. In contrast, only small child–adult differences in masking were observed for the broadband noise masker (906 components). The investigators suggested that differences in the magnitude of child–adult differences as a function of number of masker components reflected developmental effects in informational, as opposed to energetic, masking.

One explanation for the increased susceptibility to informational masking observed for children relative to adults is that children have immature sound source segregation and/or selective attention abilities. This hypothesis was tested by Lutfi et al. (2003) using a principal components analysis (PCA) to determine the number of principal components required to account for the variance across listeners in the shape of the function relating amount of masking to the number of masker components. The results indicated that a single principal component accounted for 83% of the variance both within and across age groups. This finding suggests that the large individual differences observed in studies of informational masking produced by random-frequency multitonal maskers, including differences between adults and children, are not the result of differences in detection strategy. Instead, these findings

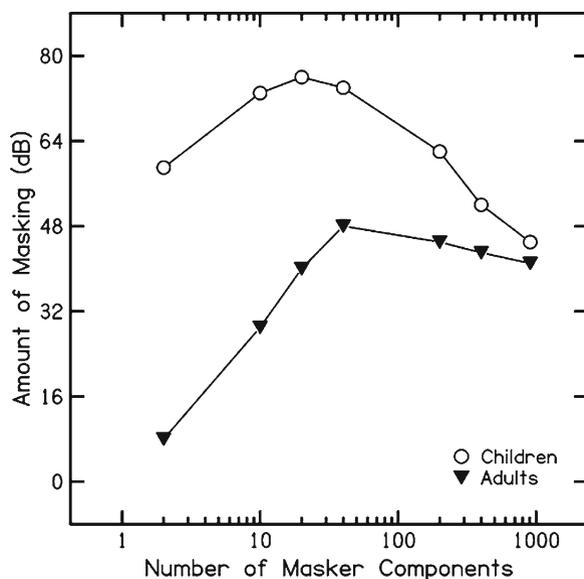


Fig. 5.4 Data from Oh et al. (2001) showing the group average amount of total masking as a function of the number of masker components for children (*open circles*) and adults (*filled triangles*). Children in this study were more susceptible to masking than adults, regardless of the number of masker components. Child–adult differences in masking were largest, however, for maskers with few components [Adapted with permission from Oh et al. (2001), Copyright © 2001 Acoustical Society of America]

support the idea that most listeners use a similar strategy, but differ in their abilities to exclusively attend to the signal frequency.

Results from the early developmental studies suggest that children have particular difficulty under conditions of high stimulus uncertainty. Recent data suggest, however, that infants and children are more susceptible to informational masking than adults in the presence of multitonal maskers, even when the masker spectra do not vary (Leibold and Werner 2006; Leibold and Neff 2007). For example, Leibold and Werner (2006) examined the ability of 7–9-month-old infants and adults to detect a 1,000-Hz pure tone presented simultaneously with a random-frequency two-tone complex, a fixed-frequency two-tone complex or a broadband noise. As shown in Fig. 5.5, infants’ thresholds were higher than adults’ in all three masker conditions, but the developmental effect was smallest in the presence of the noise expected to produce mostly energetic masking. Interestingly, a similar infant–adult difference in threshold was observed across both two-tone masker conditions. These results suggest that infants are more susceptible to informational masking compared to adults, regardless of whether the frequency components that comprise the masker are randomized or fixed. In a subsequent study, Leibold and Neff (2007) reported that school-age children (5–10 years) were also more susceptible to informational masking than adults in the presence of remote-frequency multitonal maskers with fixed

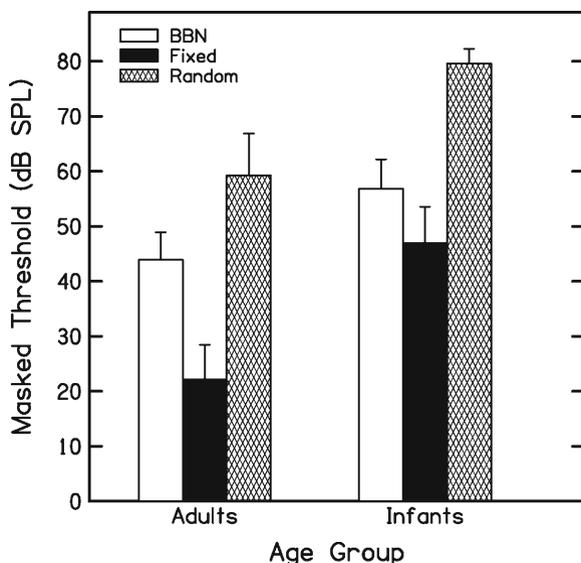


Fig. 5.5 Average masked thresholds across listeners (with SDs) are shown for infants and adults tested by Leibold and Werner (2006) for three types of maskers (*open bars* for broadband noise, *filled bars* for fixed-frequency two-tone maskers and *hatched bars* for random-frequency two-tone maskers). Infants' thresholds were higher than adults' in the presence of all three maskers. The average difference in threshold across fixed-frequency and random-frequency maskers was similar across infants and adults, however, suggesting a similar effect of masker-frequency uncertainty [Reproduced with permission from Leibold and Werner (2006). Copyright © 2006. Acoustical Society of America]

spectral components. The mechanisms responsible for these developmental effects in susceptibility to informational masking produced by multitonal maskers are not well understood, but might well reflect immature sound source segregation abilities.

Although the data are limited, a convincing argument can be made that the developmental effects in informational masking observed using multitonal stimuli extend to the perception of speech in the presence of competing background sounds. It is well documented that preschoolers and school-age children require a more advantageous signal-to-noise ratio (SNR) compared to adults to achieve similar levels of performance on speech detection or recognition tests in the presence of competing noise or speech maskers (e.g., Elliott et al. 1979; Nitttrouer and Boothroyd 1990; Hall et al. 2002; Litovsky 2005; Wightman and Kistler 2005; Panneton and Newman, Chap. 7). In addition, infants require a higher SNR than adults to recognize their own name embedded in a background of competing speech (Newman and Jusczyk 1996). Also consistent with results from the multitonal studies, child–adult differences in performance become more pronounced with increasing stimulus complexity (e.g., Hall et al. 2002; Wightman et al. 2003). Hall et al. (2002) observed larger child–adult differences for the identification of spondee words embedded in a two-talker masker compared to speech-shaped noise. Termed “perceptual masking”

(Carhart et al. 1969), the two-talker masker was expected to interfere with the target speech, in part, because listeners have difficulty performing sound source segregation. In contrast, the speech-shaped noise was expected to produce primarily energetic masking. In the adult literature, researchers have suggested that this type of speech masking is similar to the informational masking observed using multitonal stimuli (e.g., Freyman et al. 2004).

3.3 Developmental Effects in Release from Informational Masking

There is considerable debate regarding the mechanisms responsible for informational masking, but mounting evidence suggests that a failure to perceptually segregate the signal from the masker is an important contributor to the masking observed in many conditions (e.g., Kidd et al. 1994, 2002; Neff 1995; Durlach et al. 2003). Evidence supporting this hypothesis is provided by studies that have manipulated stimulus properties thought to promote sound source segregation, including spatial separation, asynchronous temporal onsets, dissimilar temporal modulations, and target-masker similarity. Substantial reductions in informational masking have been observed for most adults when these cues are provided (e.g., Kidd et al. 1994; Neff 1995; Oh and Lutfi 1998; Arbogast et al. 2002; Durlach et al. 2003).

Infants and children are more susceptible to informational masking than adults, suggesting that the ability to perceptually segregate sounds follows a prolonged time course of development. Following the adult work, investigators have started to examine the degree to which acoustic cues shown to reduce informational masking for adults improves performance for infants and school-age children (e.g., Wightman et al. 2003, 2006; Hall et al. 2005; Leibold and Neff 2007; Leibold and Bonino 2009). The results of these studies indicate that children can benefit from some, but not all, of the acoustic cues that improve performance for adults. That is, the extent to which children benefit from acoustic cues thought to aid in sound source segregation may depend on the specific cue that is manipulated (discussed by Hall et al. 2005).

Two groups of researchers have shown that children often do not benefit from lateralization cues to reduce informational masking (Wightman et al. 2003; Hall et al. 2005; but see Litovsky, Chap. 6). Wightman et al. (2003) measured detection thresholds for preschoolers (4–5 years), school-age children (6–16 years) and adults for a 1,000-Hz tone in the presence of simultaneous, random-frequency multitonal maskers with 2, 10, 20, 40, 200, 400, or 906 components. In ipsilateral conditions, the 1,000-Hz signal and multitonal masker were presented to the same ear. In contralateral conditions, the signal was presented to one ear and the masker was presented to the other ear. Masking was eliminated for most adults with the contralateral presentation. In contrast, preschoolers exhibited large amounts of masking for both ipsilateral and contralateral presentations. Although masking was reduced by an average of about 20 dB for contralateral versus ipsilateral conditions, substantial

masking and similar masking patterns were observed for preschoolers across both sets of multitonal conditions. Performance for school-age children varied widely. However, a clear age-related trend in performance was observed, with less informational masking observed for contralateral conditions with increasing age. Consistent results were reported by Hall et al. (2005) using the sequential multiple-bursts paradigm (Kidd et al. 1994). Masking release related to lateralization was assessed using a manipulation in which the signal bursts were presented to one ear and the masker bursts were presented to both ears. This presentation scheme resulted in a moderate release from informational masking for adults compared to when the signal and masker bursts were presented to the same ear. In sharp contrast, this lateralization cue did not result in masking release for any of the children tested.

Consistent with the idea that the ability to benefit from cues that promote sound segregation develops across childhood, Wightman et al. (2006) found that younger school-age children (6–9 years) did not benefit from the provision of a synchronized video of the target talker in a speech recognition task. The audiovisual cue did, however, provide a release from informational masking for both older children (12–16 years) and adults in the presence of a competing speech masker compared to presenting only the auditory information.

Several studies have observed a substantial release from informational masking in children with the introduction of temporally-based segregation cues (Hall et al. 2005; Leibold and Neff 2007; Leibold and Bonino 2009). For example, informational masking can be effectively reduced for most children when the onset of a pure-tone signal is delayed relative to the onset of a multitonal masker (Hall et al. 2005; Leibold and Neff 2007). Using a variation of the Kidd et al. (1994) multiple-bursts paradigm, Hall et al. (2005) demonstrated that a signal-masker asynchrony of 120 ms provides a similar masking release for 4–9-year-olds and adults. In the standard condition, the MBS masker consisted of two trains of eight 60-ms tone bursts. The frequency within each train of masker bursts remained constant across the eight bursts. When present, the signal was a train of eight 60-ms bursts of a 1,000-Hz pure tone played synchronously with the masker bursts. In the experimental condition, a temporal onset asynchrony of 120 ms was created between the signal and masker by increasing the number of masker bursts to 10 and delaying the onset of the eight signal bursts to begin with the third masker burst. No significant differences in masking release related to onset asynchrony cue were observed between children and adults. Similarly, Leibold and Neff (2007) reported no differences in masking release associated with a 100-ms onset/offset asynchrony in 5–10-year-olds and adults using the simultaneous multitonal paradigm (Neff and Green 1987). Results across these two studies indicate that school-age children can effectively use temporal asynchrony cues to reduce informational masking.

Recently, Leibold and Bonino (2009) reported that multiple presentations of a constant-frequency signal embedded in a random-frequency masker provide both children (5–10 years) and adults with a robust cue for reducing informational masking. Using a modified multiple-burst different paradigm (Kidd et al. 1994), detection thresholds were measured for a sequence of repeating 50-ms bursts of a 1,000-Hz pure-tone signal embedded in a sequence of 10- and 50-ms bursts of a random-frequency

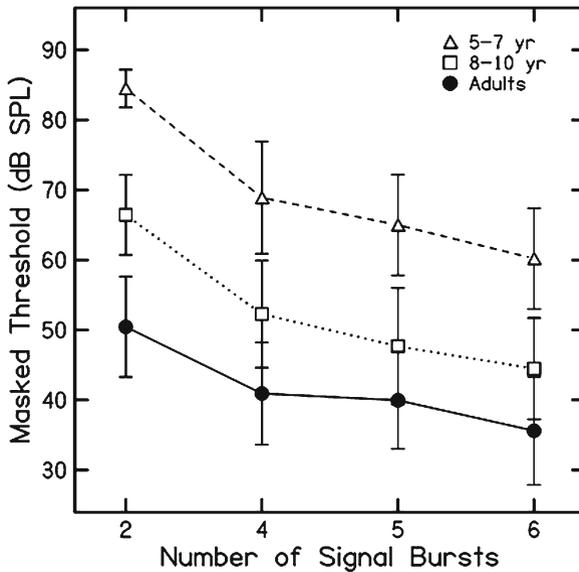


Fig. 5.6 Functions showing the relation between the number of bursts of a 1,000-Hz tone and masked thresholds in the presence of a multi-bursts different (MBD) masker reported by Leibold and Bonino (2009). Data for younger children (5–7 years) are shown by *open triangles*, data for older children (8–10 years) are shown by *open squares*, and data for adults are shown by *filled circles*. Although developmental effects in susceptibility to informational masking were observed, masked thresholds decreased by a similar amount across age as the number of signal bursts increased [Reproduced with permission from Leibold and Bonino (2009). Copyright © 2009. Acoustical Society of America]

two-tone masker. Release from informational masking was examined for conditions with two, four, five, and six signal bursts. As shown in Fig. 5.6, children were more susceptible to informational masking than adults were regardless of the number of signal bursts. Moreover, younger children (5–7 years) showed more masking than older children (8–10 years). However, masked threshold decreased with additional signal bursts by a similar amount across the three age groups. These findings are consistent with the idea that increasing the number of signal bursts aids children in the perceptual segregation of the fixed-frequency signal from the random-frequency masker, as has been previously reported for adults (Kidd et al. 2003).

3.4 Summary and Future Directions

Studies of informational masking offer a promising approach for the study of sound source segregation in infancy and childhood. These studies have consistently demonstrated that infants and children are more susceptible to informational masking than adults, providing crucial evidence regarding the potential mechanisms responsible

for developmental effects in auditory masking. Of particular importance, identifying acoustic cues that provide a release from informational masking provides valuable data regarding how the use of acoustic cues that promote the perceptual segregation of sounds changes with development. These data indicate that children can effectively use some of the same segregation cues used by adults to reduce informational masking, but other cues appear to be less salient during childhood.

Further studies are needed to determine which cues infants and children rely on to segregate sounds and whether infants and children can benefit from these same cues in more natural environments. Finally, most developmental studies have examined performance across conditions with relatively large acoustic differences. Few developmental studies have parametrically varied the strength of the grouping cue or examined the extent to which children benefit from smaller manipulations.

4 Selective Auditory Attention

After incoming acoustic waveforms have been segregated into their originating sources, the developing child must select and attend to target auditory information for further processing while ignoring components that were produced by other sound sources in their environment. This process is referred to as selective auditory attention. Given the significant role selective attention must play in the development of speech and language abilities, it is surprising that relatively few studies have examined selective attention in the auditory system. In fact, developmental research in selective attention has been devoted almost exclusively to vision. The challenge for researchers interested in the study of auditory development is that it can be difficult to determine whether attention is actually being measured rather than other processing factors or memory. Moreover, it is not always possible to separate developmental effects in parsing the auditory scene from those related to selective auditory attention. That is, a failure of selective attention can prevent infants and children from demonstrating they can segregate sounds or vice versa.

4.1 *Types of Attention*

It is important to distinguish between selective auditory attention and other aspects of attention, including arousal, orienting to sounds, sustained attention, and general inattentiveness. Arousal refers to the physiological readiness to perceive and process stimuli. An extensive review provided by Gomes et al. (2000) indicates that the arousal level of an infant changes rapidly during the first 2 months of life. This is evidenced by more time spent awake and alert as well as a larger differentiation between states (e.g., Dittrichova and Lapackova 1964). These changes in arousal level might indicate a greater readiness to perceive and select auditory objects.

Infants also appear to be selective in their orienting behavior, as evidenced by the many published studies using habituation/dishabituation paradigms. For example,

newborns respond preferentially to sounds they would have heard in utero (e.g., DeCasper and Fifer 1980; Spence and Freeman 1996). Moreover, orienting behaviors change with increasing age during infancy. Whereas novel stimuli often elicit an orienting response in early infancy, the orienting response becomes smaller and higher cognitive processes appear to have a greater effect on orienting with increasing age (e.g., Thompson and Weber 1974; Morrongiello and Clifton 1984).

One type of attention that has received considerable study is sustained attention, often referred to as “inattentiveness” in the developmental psychoacoustics literature. General inattentiveness has been considered as an explanation for the prolonged time course of development observed for many behavioral measures of hearing. That is, infants and children may perform more poorly than adults because they are “off-task” and fail to obtain any information about the target sound on some proportion of trials. For example, general inattentiveness has been proposed as a model to explain age differences in auditory sensitivity. As described earlier in this volume (Buss, Hall, and Grose, Chap. 4), infants’ absolute thresholds are higher than adults’. It has been suggested that these findings reflect that infants are less attentive and “guess” on a larger proportion of trials compared to adults. This model is supported by observations that infants’ and young childrens’ psychometric functions for detection often asymptote at a value less than 100% correct (Allen and Wightman 1994; Bargones et al. 1995; Werner and Boike 2001). However, several investigators have examined the impact of inattentiveness on the psychometric function and have determined that it cannot explain the full threshold difference (e.g., Viemeister and Schlauch 1992; Wightman and Allen 1992; Werner and Boike 2001). For example, Werner and Boike (2001) demonstrated that models of inattentiveness can account for no more than 2–3 dB of the difference in detection threshold between infants and adults. Nonetheless, infants and young children show reduced performance levels compared to adults, even for stimuli that are clearly audible. The psychometric function continues to mature throughout the preschool years (Werner and Marean 1996), indicating that the ability to sustain attention improves with increasing age.

4.2 *Selective Auditory Attention*

Cherry (1953) first described the “cocktail party problem” as the remarkable ability of listeners to selectively attend to a single voice in the presence of many. Early adult studies of selective attention used variations of Cherry’s shadowing paradigm (e.g., Cherry 1953; Broadbent 1958; Treisman 1960). In the shadowing paradigm, two messages are played simultaneously over headphones, a target and a distracter message. Listeners are asked to shadow the target message by repeating it back without delay while ignoring the distracter message. Cherry (1953) observed that adults could shadow the target message easily if the target and distracter messages were presented dichotically. Moreover, adults rarely noticed when changes were made to the message in the unattended ear, including a switch to a foreign language or reversing the speech. Listeners did, however, notice a change in the unattended ear from speech to a pure

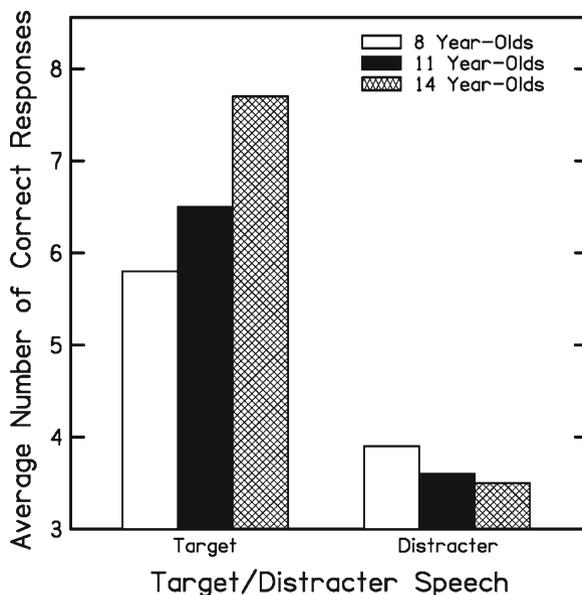


Fig. 5.7 Data from Doyle (1973) showing the mean number of correct target and distracter responses for children ages 8, 11, and 14 years asked to attend to a target speech message and ignore a different speech message. Age-related improvements were observed in the ability to shadow the target speech. In addition, the number of intrusions from the distracter message decreased with age [Adapted with permission from Doyle (1973). Copyright © 1973. Elsevier]

tone or from a male to a female voice. Cherry interpreted these results as evidence that only simple physical characteristics of sound were processed in the unattended ear.

Shadowing paradigms have been used to examine selective attention in children (e.g., Maccoby 1969; Doyle 1973; Wightman and Kistler 2005). In general, children have more difficulty shadowing the target message compared to adults and their responses often reflect multiple intrusions from the distracting speech message. In addition, older children perform better and make fewer errors than younger children (e.g., Maccoby 1969; Doyle 1973). This developmental trend suggests that selective auditory attention develops throughout childhood. In one study, Doyle (1973) asked children ages 8, 11, and 14 years to listen only to target speech while ignoring a distracting speech message presented to the same ear. As shown in Fig. 5.7, an age-related improvement was observed in children's ability to report the target speech correctly. Younger children were also more likely to report words or partial words from the distracting voice compared to older children. Based on these findings, Doyle (1973) suggested that younger children are less able to disregard irrelevant auditory information compared to older children.

Consistent results were reported in a more recent study by Wightman and Kistler (2005), who examined auditory selective attention in children (4–16 years) and adults using a closed-set speech recognition task. In one condition, listeners were presented with both a target and a distracter speech message in the same ear. In the remaining

two conditions, an additional distracter speech message or a speech-shaped noise was presented to the opposite ear. Similar to the early shadowing studies, younger children performed more poorly than older children. Interestingly, some of the oldest children tested continued to perform more poorly than adults. These results suggest that the development of selective auditory attention extends into adolescence.

In addition to the behavioral data, shadowing studies using event-related potentials (ERPs) indicate that the central pathways responsible for selective auditory attention are not mature until after 8 years of age. Coch et al. (2005) recorded event-related potentials (ERPs) in the context of a dichotic listening paradigm. Children (6–8 years) and adults were asked to attend to one of two narratives presented at the same time to different ears. The N1 attention effect previously observed for adults (Hillyard et al. 1973) was not adult-like for even the oldest children tested.

Additional evidence that selective auditory attention continues to develop into childhood comes from studies examining age-related changes in performance on psychoacoustic tasks across infancy and childhood. Results from these studies suggest that infants and children often listen less selectively than adults while performing these tasks (e.g., Werner and Bargones 1991; Allen and Wightman 1994; Bargones and Werner 1994; Oh et al. 2001). For example, adults listen selectively for an expected stimulus when detecting a tone in noise. This frequency-selective listening strategy improves adults' performance for detecting the expected frequency, but at the expense of detection performance at unexpected frequencies (e.g., Dai et al. 1991). Unlike adults, 7–9-month-old infants do not appear to listen selectively in the frequency domain during detection tasks. In their classic study, Bargones and Werner (1994) measured detection thresholds for pure tones in the presence of continuous broadband noise. In a mixed condition, one of three possible signal frequencies was presented on each trial. The expected signal frequency was presented on the majority of trials (75%) whereas unexpected probe frequencies were presented on the remaining trials (25%). Performance for both unexpected and expected tones was compared to performance for a fixed condition in which only one frequency was presented. As in previous studies (e.g., Dai et al. 1991), adults' performance was better for expected versus unexpected frequencies. In contrast, infants' performance for the unexpected frequencies was similar to their performance for the expected frequency. One interpretation for these results is that adults listen selectively in the frequency domain, but infants monitor a broad range of frequencies.

Consistent with the hypothesis that infants have difficulty focusing their attention in the frequency domain, infants show masking in conditions associated with little or no masking for adults (e.g., Werner and Bargones 1991; Leibold and Werner 2006). Werner and Bargones (1991) observed that 6-month-old infants are susceptible to remote-frequency or “distraction” masking, in that a remote-frequency noise band (4,000–10,000 Hz) can produce significant simultaneous masking of a 1,000-kHz signal. Group average thresholds for infants were approximately 10 dB higher in the presence of the remote-frequency noise compared to quiet, regardless of whether the overall level of the masker was 40 or 50 dB SPL. In contrast, similar thresholds were observed for adults across quiet and remote-frequency noise conditions. Observing roughly the same thresholds with a 10-dB increase in masker level is

inconsistent with energetic (peripheral) masking effects (e.g., Moore et al. 1997). Instead, this result indicates more central effects such as immature selective auditory attention. Recently, Leibold and Neff (2011) observed smaller, but significant, remote-frequency masking effects in 4–6-year-old children. Using the general stimuli and approach described by Werner and Bargones (1991) to test infants, no systematic masking effects were observed for 7–9-year-olds or adults, but average thresholds in noise were elevated 3.5 dB for 4–6-year-olds. In addition, no growth of masking was observed for 40- versus 60-dB maskers, suggesting nonperipheral masking effects for the youngest children.

Data describing the relation between selective auditory attention and performance on conventional psychoacoustic tasks in preschoolers and older children remain limited, but are consistent with the idea that at least young school-age children have difficulty focusing their attention on the most important components of complex sounds. That is, they may give undue perceptual weight to irrelevant information. For example, Stellmack et al. (1997) reported child–adult differences in perceptual weighting strategies for a selective listening task. Five-year-old children and adults were asked to attend to a target tone while ignoring two “distractor” tones in the context of an intensity discrimination procedure. No differences in perceptual weights were observed between children and adults when the levels of the distractor tones were relatively low, with most listeners assigning the greatest weight to the target component. Child–adult differences in weights were observed; however, as the intensity levels of the distractor tones were increased. Adults continued to assign the greatest weight to the target tone, but children tended to assign the same weight to the target and two distractor tones. This pattern of results suggests that young school-age children listen less selectively than adults, but only when the intensity level of the distracting information is relatively high.

Finally, it has been suggested that infants listen in a broad way to learn the important cues of speech across a variety of different listening contexts (Werner 2007). Studies of children’s speech perception support this hypothesis. For example, 4-year-olds are more influenced by global and dynamic speech cues such as formant transitions when they are asked to categorize fricatives. In contrast, 7-year-olds and adults tend to rely on more detailed cues such as the spectra of the noise (reviewed by Nittrouer 2006). These results are consistent with the idea that young children listen to speech more broadly than do older children or adults, a strategy that might enable children to learn the important features of speech. Panneton and Newman (Chap. 7) provide a complete discussion of the development of speech perception, including the influence of remote-frequency masking on measures of speech perception.

4.3 Summary and Future Directions

Selective attention to sound appears to play a major role in the development of hearing (e.g., Doyle 1973; Wightman and Kistler 2005). Adults, as experienced listeners, can listen selectively at a specific point in time, to a specific feature of speech, to one

location in space, or to a particular frequency within a complex sound (reviewed by Hafter et al. 2008). In contrast, developmental studies using both speech and non-speech stimuli have provided evidence that infants and young children listen less selectively than adults. As a consequence, infants and children tend to listen to all of the information in a sound rather than focusing on the most important features (reviewed by Werner 2007). An important consequence of this broadband listening strategy is that it can result in masking from background sounds that produce little or no masking for adults (e.g., Werner and Bargones 1991; Leibold and Neff 2007).

During infancy and childhood, performance improves with age on tasks that require the listener to actively direct attention, although the source of this improvement is unclear. One challenge for researchers interested in the development of selective auditory attention is that infants and children may have more difficulty segregating sound sources compared to adults (discussed in Sect. 2). If this is true, what we are observing may reflect an inability to segregate the sound source rather than an inability to attend to it selectively. Future investigations require the use of stimuli that are known to be easily segregated by listeners of all ages.

5 Summary

The peripheral auditory system appears to provide the brain with an accurate representation of sound by about 6 months of age, but human auditory development is a prolonged process. This development requires increasing sophistication in the skills needed to separate and select target sounds in complex acoustic environments. What infants and children hear when they listen to complex sounds is different than what adults hear.

As discussed at the end of each section, there are numerous unresolved issues to be addressed by future studies of auditory scene analysis and auditory attention. For example, a fundamental question for both researchers and clinicians is whether the amount of quality of early auditory experiences influences the development of skills related to auditory scene analysis and selective attention. The time course for typical auditory development might be altered in children with hearing loss if their access to sound is degraded or reduced.

Acknowledgments This work was supported by NIH grant R03 DC00838.

References

- Alexander, J. M., & Lutfi, R. A. (2004). Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights. *Journal of the Acoustical Society of America*, 116, 2234–2247.
- Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research*, 37, 205–215.

- Allen, P., & Wightman, F. (1995). Effects of signal and masker uncertainty on children's detection. *Journal of Speech and Hearing Research*, 38, 503–511.
- Arbogast, T. L., Mason, C. R., & Kidd, G. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, 112, 2086–2098.
- Bargones, J. Y., & Werner, L. A. (1994). Adults listen selectively; infants do not. *Psychological Science*, 5, 170–174.
- Bargones, J. Y., Werner, L. A., & Marean, G. C. (1995). Infant psychometric functions for detection: Mechanisms of immature sensitivity. *Journal of the Acoustical Society of America*, 98, 99–111.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in sounds* (pp. 10–36). New York: Oxford University Press.
- Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244–249.
- Bregman, A. S., & Pinker, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 32, 19–31.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.
- Broadbent, D. E., & Ladefoged, P. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 31, 1539–1539.
- Brungart, D. D. (2005). Informational and energetic masking effects in multitalker speech perception. In P. Divenyi (Ed.), *Speech separation in humans and machines* (pp. 261–267). New York: Springer.
- Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America*, 45, 694–703.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25, 975–979.
- Coch, D., Sanders, L. D., & Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *Journal of Cognitive Neuroscience*, 17, 605–622.
- Dai, H., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *Journal of the Acoustical Society of America*, 89, 2837–2842.
- Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (pp. 387–424). New York: Academic Press.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174–1176.
- Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development*, 5, 261–276.
- Dittrichova, J., & Lapackova, V. (1964). Development of the waking state in young infants. *Child Development*, 35, 365–370.
- Doyle, A. B. (1973). Listening to distraction: A developmental study of selective attention. *Journal of Experimental Child Psychology*, 15, 100–115.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*, 114, 368–379.
- Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, H. S., & Kidd, G. (2005). Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers. *Journal of the Acoustical Society of America*, 118, 2482–2497.
- Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., & Katz, D. (1979). Children's understanding of monosyllabic nouns in quiet and noise. *Journal of the Acoustical Society of America*, 66, 12–21.
- Fassbender, C. (1993). *Auditory grouping and segregation processes in infancy*. Doctoral dissertation. Norderstedt, Germany: Kaste Verlag.
- Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, 12, 47–65.
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America*, 115, 2246–2256.

- Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., & Cowan, N. (2000). The development of auditory attention in children. *Frontiers in Bioscience*, 5, 108–120.
- Grose, J. H., Hall, J. W., & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech and Hearing Research*, 36, 351–356.
- Hafer, E. R., Sarampalis, A., & Loui, P. (2008). Auditory attention and filters. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory perception of sound sources* (pp. 115–142). New York: Springer.
- Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, 76, 50–56.
- Hall, J. W., Grose, J. H., & Dev, M. B. (1997). Auditory development in complex tasks of comodulation masking release. *Journal of Speech, Language, and Hearing Research*, 40, 946–954.
- Hall, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, 23, 159–165.
- Hall, J. W., Buss, E., & Grose, J. H. (2005). Informational masking release in children and adults. *Journal of the Acoustical Society of America*, 118, 1605–1613.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, 182, 177–180.
- Huang, R., & Richards, V. M. (2006). Coherence detection: Effects of frequency, frequency uncertainty, and onset/offset delays. *Journal of the Acoustical Society of America*, 119, 2298–2304.
- Kidd, G., Mason, C. R., Deliwala, P. S., Woods, W. S., & Colburn, H. S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, 95, 3475–3480.
- Kidd, G., Mason, C. R., & Dai, H. (1995). Discriminating coherence in spectro-temporal patterns. *Journal of the Acoustical Society of America*, 97, 3782–3790.
- Kidd, G., Mason, C. R., & Arbogast, T. L. (2002). Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*, 111, 1367–1376.
- Kidd, G., Mason, C. R., & Richards, V. M. (2003). Multiple bursts, multiple looks, and stream coherence in the release from informational masking. *Journal of the Acoustical Society of America*, 114, 2835–2845.
- Leibold, L., & Neff, D. L. (2007). Effects of masker-spectral variability and masker fringes in children and adults. *Journal of the Acoustical Society of America*, 121, 3666–3676.
- Leibold, L., & Neff, D. L. (2011). Masking by a remote-frequency noise band in children and adults. *Ear and Hearing*, doi: 10.1097/AUD.0b013e31820e5074.
- Leibold, L. J., & Bonino, A. Y. (2009). Release from informational masking in children: Effect of multiple signal bursts. *Journal of the Acoustical Society of America*, 125, 2200–2208.
- Leibold, L. J., & Werner, L. A. (2006). Effect of masker-frequency variability on the detection performance of infants and adults. *Journal of the Acoustical Society of America*, 119, 3960–3970.
- Leibold, L. J., Hitchens, J. J., Neff, D. L., & Buss, E. (2010). Excitation-based and informational masking of a tonal signal in a four tone masker. *Journal of the Acoustical Society of America*, 127, 2441–2450.
- Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*, 117, 3091–3099.
- Lutfi, R. A., Kistler, D. J., Oh, E. L., Wightman, F. L., & Callaghan, M. R. (2003). One factor underlies individual differences in auditory informational masking within and across age groups. *Perception and Psychophysics*, 65, 396–406.
- Maccoby, E. E. (1969). The development of stimulus selection. *Minnesota Symposia on Child Psychology*, 3, 68–96.
- McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal of the Acoustical Society of America*, 102, 2945–2953.
- Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45, 224–240.
- Morronegiello, B. A., & Clifton, R. K. (1984). Effects of sound frequency on behavioral and cardiac orienting in newborn and five-month-old infants. *Journal of Experimental Child Psychology*, 38, 429–446.

- Neff, D. L. (1995). Signal properties that reduce masking by simultaneous, random-frequency maskers. *Journal of the Acoustical Society of America*, 98, 1909–1920.
- Neff, D. L., & Callaghan, B. P. (1988). Effective properties of multicomponent simultaneous maskers under conditions of uncertainty. *Journal of the Acoustical Society of America*, 83, 1833–1838.
- Neff, D. L., & Dethlefs, T. M. (1995). Individual differences in simultaneous masking with random-frequency, multicomponent maskers. *Journal of the Acoustical Society of America*, 98, 125–134.
- Neff, D. L., & Green, D. M. (1987). Masking produced by spectral uncertainty with multicomponent maskers. *Perception and Psychophysics*, 41, 409–415.
- Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception and Psychophysics*, 58, 1145–1156.
- Nittrouer, S. (2006). Children hear the forest. *Journal of the Acoustical Society of America*, 120, 1799–1802.
- Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*, 87, 2705–2715.
- Oh, E. L., & Lutfi, R. A. (1998). Nonmonotonicity of informational masking. *Journal of the Acoustical Society of America*, 104, 3489–3499.
- Oh, E. L., Wightman, F., & Lutfi, R. A. (2001). Children's detection of pure-tone signals with random multitone maskers. *Journal of the Acoustical Society of America*, 109, 2888–2895.
- Richards, V. M., & Neff, D. L. (2004). Cuing effects for informational masking. *Journal of the Acoustical Society of America*, 115, 289–300.
- Richards, V. M., Tang, Z., & Kidd, G. D. (2002). Informational masking with small set sizes. *Journal of the Acoustical Society of America*, 111, 1359–1366.
- Spence, M. J., & Freeman, M. S. (1996). Newborn infants prefer the maternal low-passed filtered voice, but not the maternal whispered voice. *Infant Behavior and Development*, 19, 199–212.
- Stellmack, M. A., Willighnganz, M. S., Wightman, F. L., & Lutfi, R. A. (1997). Spectral weights in level discrimination by preschool children: Analytic listening conditions. *Journal of the Acoustical Society of America*, 101, 2811–2821.
- Sussman, E., Wong, R., Horvath, J., Winkler, I., & Wang, W. (2007). The development of the perceptual organization of sound by frequency separation in 5–11-year-old children. *Hearing Research*, 225, 117–127.
- Thompson, G., & Weber, B. A. (1974). Responses of infants and young children to behavior observation audiometry (BOA). *Journal of Speech and Hearing Research*, 39, 140–147.
- Treisman, A. M. (1960). Contextual cues in selective listening. *The Quarterly Journal of Experimental Psychology*, 12, 242–248.
- van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Doctoral dissertation. Eindhoven, Holland: Technische Hogeschool.
- Veloso, K., Hall, J. W., & Grose, J. H. (1990). Frequency selectivity and comodulation masking release in adults and in 6-year-old children. *Journal of Speech and Hearing Research*, 33, 96–102.
- Viemeister, N. F., & Schlauch, R. S. (1992). Issues in infant psychoacoustics. In L. A. Werner, & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 191–209). Washington, DC: American Psychological Association.
- Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science*, 164, 586–587.
- Watson, C. S., Wroton, H. W., Kelly, W. J., & Benbassat, C. A. (1975). Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent interval. *Journal of the Acoustical Society of America*, 57, 1175–1185.
- Werner, L. A. (2007). What do children hear: How auditory maturation affects speech perception. *The ASHA Leader*, 12, 6–7, 32–33.
- Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception and Psychophysics*, 50, 405–412.

- Werner, L. A., & Boike, K. (2001). Infants' sensitivity to broadband noise. *Journal of the Acoustical Society of America*, 109, 2103–2111.
- Werner, L. A., & Marean, C. G. (1996). *Human auditory development*. Boulder: Westview Press.
- Wightman, F., & Allen, P. (1992). Individual differences in auditory capability among preschool children. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 113–133). Washington, DC: American Psychological Association.
- Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*, 118, 3164–3176.
- Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., & Oh, E. (2003). Children's detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*, 113, 3297–3305.
- Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *Journal of the Acoustical Society of America*, 119, 3940–3949.
- Wright, B. A., & Saberi, K. (1999). Strategies used to detect auditory signals in small sets of random maskers. *Journal of the Acoustical Society of America*, 105, 1765–1775.
- Yost, W. (1997). The cocktail party problem: Forty years later. In R. H. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 329–348). Hillsdale, NJ: Lawrence Erlbaum.
- Zettler, C. M., Sevcik, R. A., Morris, R. D., & Clarkson, M. G. (2008). Comodulation masking release (CMR) in children and the influence of reading status. *Journal of Speech, Language, and Hearing Research*, 51, 772–784.

Chapter 6

Development of Binaural and Spatial Hearing

Ruth Y. Litovsky

1 Background on Binaural Hearing

When a person hears sounds in the environment, there are several important tasks that the auditory system must accomplish, such as determining the location of sound sources and the meaning of those sources. These tasks are relevant to children who spend time every day in noisy environments, such as classrooms, and adults who have to operate in complex auditory environments. The auditory mechanisms that enable listeners to accomplish these tasks are generally thought to involve binaural processing. Acoustic information arriving at the two ears is compared at the level of the brain stem, combined and transmitted to the central auditory system for further analysis. As discussed in the text that follows, in listeners with hearing loss these mechanisms may be compromised or not fully developed.

This chapter focuses on the development of perceptual abilities that depend on binaural and spatial hearing. In particular, there is a focus on psychophysics research in children and infants, conducted since the review by Litovsky and Ashmead (1997). A well known issue in developmental psychoacoustics is the ability of the investigator to measure reliable changes in the behavior of the young listener. Toward that goal, a number of behavioral measures have been implemented that are first discussed in the chapter, laying a framework for the psychophysical results. Finally, translational issues require careful consideration of the potential utility of binaural hearing in populations of young, hearing-impaired, and deaf children. This topic is covered later in the chapter, with an eye toward identifying key issues that remain unresolved, as well as gleaning from this research any preliminary conclusions about the role of auditory plasticity in the emergence of binaural and spatial hearing skills.

R.Y. Litovsky (✉)
University of Wisconsin Waisman Center, Madison, WI, USA
e-mail: Litovsky@waisman.wisc.edu

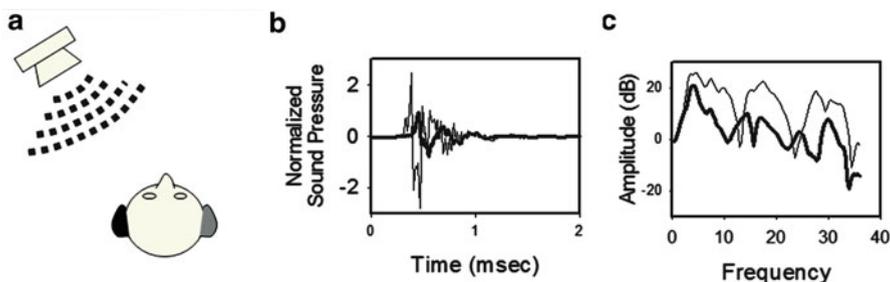


Fig. 6.1 In (a), a schematic of a sound source presented at 45° to the *left* of the listener is depicted. (b, *left*) The time waveforms of impulse responses recorded in the *left* and *right* ear canals, for that sound source. (b) The amplitude spectra for the same source, also recorded in the *left* and *right* ear canals. The *left* ear is shown in *thick* lines and the *right* ear in *thin* lines. The *left* ear response occurs sooner than the *right* ear response (see c), hence the interaural time difference (ITD). In addition, the *left*-ear response has a greater amplitude (see b), hence the interaural level difference (ILD)

1.1 Auditory Cues that Are Potentially Available When Two Ears Are Stimulated

It is important to distinguish between listening situations in which one ear is stimulated (monaural) as opposed to those in which two ears are stimulated (bilateral or binaural). In a normally functioning auditory system, when sounds reach the ears from a particular location in space, the spherical shape of the head renders an important set of acoustic cues. In the horizontal plane, sources presented from directly in front or behind reach the ears at the same time and with the same intensity. Sources that are displaced to the side will reach the near ear before reaching the far ear. Thus, a binaural cue known as interaural time difference (ITD) varies with spatial location; however, the auditory system is particularly sensitive to ITD at frequencies below 1,500 Hz. For amplitude-modulated signals such as speech, ITD cues are also available from differences in the timing of the envelopes (slowly varying amplitude) of the stimuli. Interaural level difference (ILD) is a second binaural cue that results from the fact that the head creates an acoustic “shadow,” so that the near ear receives a greater intensity than the far ear. ILDs are particularly robust at high frequencies and can be negligible at frequencies as low as 500 Hz. In addition to these binaural cues, listeners have access to monaural cues: directionally dependent spectral cues introduced by reflections from the pinnae, head, and torso. These are particularly relevant when sources are displaced in elevation (Middlebrooks and Green 1991). Figure 6.1 provides a schematic of the directionally dependent cues that would be potentially available to listeners in the horizontal plane for a brief signal such a click.

1.2 Binaural Hearing Phenomena in Adult Listeners

1.2.1 Sound Localization

The availability of binaural cues enables listeners to function on numerous tasks that involve listening in realistic, complex auditory environments, beginning with localizing sound sources. Sound localization tasks reflect the extent to which an organism has a well developed spatial-hearing “map,” a perceptual representation of where sounds are located relative to the head. This ability has been studied in adult listeners for decades, with a focus on the extent to which listeners can either identify the exact position of a sound or discriminate changes in the location of sounds. In the former case, localization accuracy has been measured by asking listeners to point to source locations, either by using a pointer object, directing their nose or head in that direction, or labeling source locations using numerical values that refer to angles in the horizontal and/or vertical dimension. Accuracy depends on factors such as the task, instructions to listeners, stimulus frequency content and duration, and response options. Error rates are often quantified by the root-mean-square (RMS) error, an estimate of the deviation of responses from the true source location. Results from some studies suggest that RMS errors can be as small as a few degrees (Hartmann 1983; Middlebrooks and Green 1991; Populin 2008). The relevance to children, as discussed later, is that responses regarding perceived source location are challenging to obtain from children, and thus, very few interpretable, developmental studies of localization exist.

1.2.2 Sound Source Discrimination

In contrast with absolute localization measures, the ability of children to discriminate between stimuli whose locations vary has been studied more extensively. In free field, a common measure of discrimination thresholds is the minimum audible angle (MAA), the smallest change in the angular position of a sound source that can be reliably discriminated (Mills 1958; Litovsky and Macmillan 1994). In addition, stimulus manipulations, such as restricting the bandwidth, can lead to poorer MAA (Mills 1958). Although MAA thresholds have been measured in adults in numerous studies aiming to evaluate the limits of the auditory system, the relation between MAA and absolute localization has not been clearly established. That is, the ability of a listener to discriminate small changes in source location does not automatically generalize to having acute sound localization abilities (Hartmann and Rakerd 1989). As discussed later, the MAA task has been implemented in developmental research because the paradigm is easily adapted to listeners who are not able to identify source locations reliably, such as infants and children. Because of the nature of the task only requires discrimination on each trial, infants can be trained to respond selectively between two options.

1.2.3 Binaural Unmasking

In complex auditory environments, multiple sounds occur that vary in content and direction. The ability to segregate target speech from competing speech or noise is determined by a complex set of auditory computations that involve both monaural and binaural processes and that depend on features of the competing sounds (Hawley et al. 1999, 2004; Bronkhorst 2000; Culling et al. 2004). Spatial cues play a key role in facilitating source segregation: speech intelligibility improves when the target speech and competing sounds are spatially separated, resulting in spatial release from masking (SRM) (Plomp and Mimpen 1981; Hawley et al. 1999, 2004; Arbogast et al. 2002; Bronkhorst 2000). SRM can be quite large (10–12 dB) or relatively small (3–5 dB). The larger effects seem to occur when the competing sound, or “masker,” and target can be easily confused, and when listeners are unsure as to what aspects of the masker to ignore (i.e., when “informational masking” occurs; see Leibold, Chap. 5). Spatial separation of maskers from the target is an effective way to counteract informational masking (Kidd et al. 1998; Freyman et al. 1999; Arbogast et al. 2002). As a result, the magnitude of SRM with speech-on-speech masking can be quite large relative to noise maskers (Durlach et al. 2003; Jones and Litovsky 2008). The magnitude of SRM can also be divided into binaural and monaural components (Hawley et al. 2004). Binaural-dependent SRM occurs when target and maskers differ in ITDs and/or ILDs, whereas SRM can also occur under condition in which the monaural “head shadow” occurs.

The binaural cues that are thought to be important for segregation of speech and noise can be studied selectively over headphones by imposing either similar binaural cues on the speech and masker, the collocated condition, or by varying the binaural cues such that the target and masker are perceived to be at different intracranial locations, the separated condition. For speech separation, the binaural intelligibility level difference (BILD), the difference in speech intelligibility threshold between the collocated and separated conditions, can be as large as 12 dB in adults, depending on the condition (Blauert 1997; Hawley et al. 2004). A simpler version of the BILD is the binaural masking level difference (BMLD), where a target signal such as a tone or narrow-band noise is detected in the presence of a masking noise. BMLD can be measured, for example, by comparing threshold for tone detection when: Both the noise and tone are in-phase at the two ears, the N_0S_0 condition, and when the noise is in-phase at the two ears but the tone signal is out-of-phase at the two ears, the N_0S_π condition. Presumably, the tone and noise are perceived as collocated intracranially in the N_0S_0 condition, while they are perceived as spatially separated in the N_0S_π condition. The difference in threshold between these conditions ranges from 8 to 30 dB, depending on the specific condition (Zurek and Durlach 1987; van de Par and Kohlrausch 1999).

The BMLD, headphone-based paradigm, in which ITD is manipulated to produce source segregation, has been instructive in considering the benefit that listeners gain in spatially separated conditions in free field. Unmasking occurs in these paradigms because in the “separated” conditions the acoustic characteristics of the signals in two ears are highly dissimilar (Gabriel and Colburn 1981; Bernstein and Trahiotis 1992; Culling and Summerfield 1995). Thus the task becomes one in which listeners

detect “incoherence” between the separated and colocated conditions. While this area of research is extremely important in the field of binaural psychoacoustics, few studies exist in which the development of BMLD has been investigated (see review of this work in Sect. 3.5). This area provides fertile ground for future research.

1.2.4 The Precedence Effect

In reverberant environments, sound arrives at the listener’s ears through a direct path, which is the most rapid and least disturbed path. Reflections of the sound from nearby surfaces, including walls and various objects, reach the ears with a time delay, and offer their own set of localization cues. In reverberant rooms, although listeners are aware of the presence of reflections, localization cues carried by reflections are deemphasized relative to the cues carried by the source. This phenomenon is commonly attributed to auditory mechanisms that assign greater weight to the localization cues belonging to the preceding, or first-arriving sound, hence referred to as the precedence effect (PE). For reviews see Blauert (1997) and Litovsky et al. (1999). Although the PE is not a direct measure of people’s ability to function in realistic acoustic environments, it provides an excellent tool for exploring how the auditory system assigns greater weight to the location associated with the source.

Studies on the PE typically utilize simplified stimulus paradigms in which one source (lead) is presented from a given location, or bears a set of binaural cues under headphones, and an unrealistic reflection (lag) is simulated, whose intensity is typically the same as that of the lead. The stimulus feature that is generally varied in PE studies is the time delay between the onsets of the lead and lag, as illustrated in Fig. 6.2a. At short delays the lead and lag perceptually fuse into a single auditory percept; when the delay is between 0 and 1 ms, summing localization occurs, whereby both lead and lag contribute to the perceived location of the fused image. As the delay increases to 1 ms and beyond, the location of the lead dominates the perceived location of the fused auditory image, a phenomenon that has become known as localization dominance (Litovsky et al. 1999). The delay at which the lead and lag break apart into two auditory events is known as fusion echo threshold. Another way of quantifying the extent to which the directional cues from the lag are available to the listener is to measure discrimination suppression, whereby the listener discriminates changes in the location, or interaural parameters related to the lag. As delays increase, the ability of the listener to extract directional cues from the lag improves, indicating that discrimination suppression diminishes with delay in a fashion similar to reduced fusion.

2 Behavioral Measures

While objective measures, such as imaging techniques (see Eggermont and Moore, Chap. 3) are used to understand the neural basis of auditory function, behavioral measures allow one to evaluate the perceptual consequences of auditory signals on the listener. However, the greatest challenge to obtaining valid and interpretable

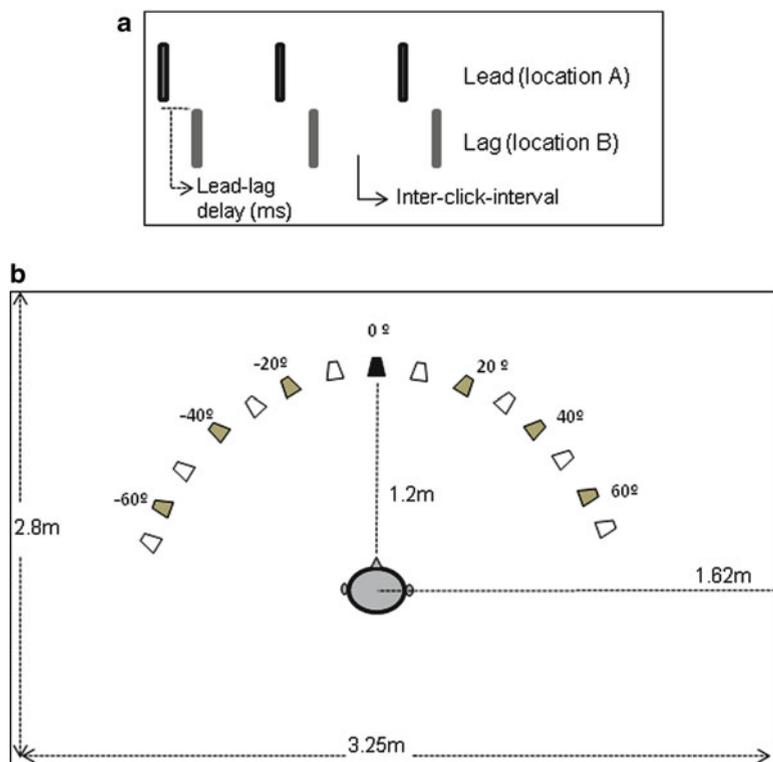


Fig. 6.2 Stimulus configurations used in free field experiments on the precedence effect. (a) Schematic diagram of the temporal order of the stimuli, which consisted of lead and lag sources, presented at a defined interclick interval (i.e., the interval between the onsets of two successive lead sources). In addition, the lead-lag delay denotes the onset delay between the lead and lag in a given click pair. (b) The spatial locations of the lead and lag in a sound field when localization dominance or fusion were measured [b from Litovsky and Godar (2010)]

results is gaining assurance that the methods used for obtaining data from behaving infants and children are valid and interpretable.

Source location discrimination, a measure of spatial hearing, has been measured successfully (e.g., Ashmead et al. 1986; Litovsky 1997) by building on the clinically based visual reinforcement audiometry (Moore et al. 1975). Although infants can display excellent right–left discrimination at relatively small angles, this task cannot provide information regarding spatial hearing, that is, *where* the sound source is perceived. Similarly, the observer-based Psychophysical Procedure (OPP) was developed by Werner and colleagues to assess auditory sensitivity in young infants using behaviors that are not strictly restricted to overt head turning (Olsho et al. 1987, 1988; Spetner and Olsho 1990), in particular because young infants do not turn their heads reliably until 5–6 months of age. The OPP was initially used to assess auditory abilities such as pure tone sensitivity and frequency discrimination

(Olsho et al. 1987; Spetner and Olsho 1990), forward masking (Werner 1999), and informational masking (Leibold and Werner 2006). More recently, the OPP has been implemented with older listeners (e.g., toddlers age 2.5 years) and used to study spatial hearing on a right–left discrimination task for sounds that were well above audibility threshold (Grieco-Calub et al. 2008). In this situation, the OPP was useful because, although these young children have the capacity to emit overt head turn, they responded consistently to sound source location with more subtle behaviors that were captured with the OPP.

Although the head turning and observer-based methods have the unique advantage of enabling testing of very young listeners, they are limited to measures of discrimination. To evaluate more real-world aspects of spatial hearing abilities, one must employ measures in which multiple response options are possible. A unique set of studies on the emergence of spatial hearing skills in young infants was initiated by Clifton and colleagues (Perris and Clifton 1988; Clifton et al. 1991; Litovsky and Clifton 1992), whereby localization accuracy was evaluated using a behavior that infants use in their everyday interactions, that of reaching. By 5–6 months of age, infants reach proficiently, using visual cues to guide their motor behaviors. Clifton and colleagues demonstrated that the reaching behavior can be elicited in response to sound, in the absence of visual cues, such that infants will readily reach for sounding objects in the dark.

The methods used in the aforementioned experiments that have been implemented with young listeners were developed for the purpose of eliciting responses from listeners who cannot be instructed regarding the task. Thus, one can measure changes in behavior that are interpreted as indicative of changes in sensation and perception. Once children are old enough to engage in verbal discussion, to read, and to communicate with the investigator, testing can take the shape of psychophysical testing approaches that are commonly used with adults. However, incentives for participation are provided, and interactive computerized platforms have been developed in a number of laboratories as a means of providing ongoing feedback and reinforcement to the children. The studies that are described in the text that follows relied on these methods to elicit responses with children ages 3 years and older.

3 Psychophysical Studies on the Development of Spatial Hearing

3.1 Developmental Trends in Spatial Hearing Acuity

When measuring spatial hearing in children, the question of interest is typically: How well can a child identify where a sound is coming from? For children with normal hearing, this ability is present in a rudimentary form at birth. Newborn infants orient toward the direction of auditory stimuli within hours after birth. The early head-orienting response is an unconditioned reflex that enables the infant

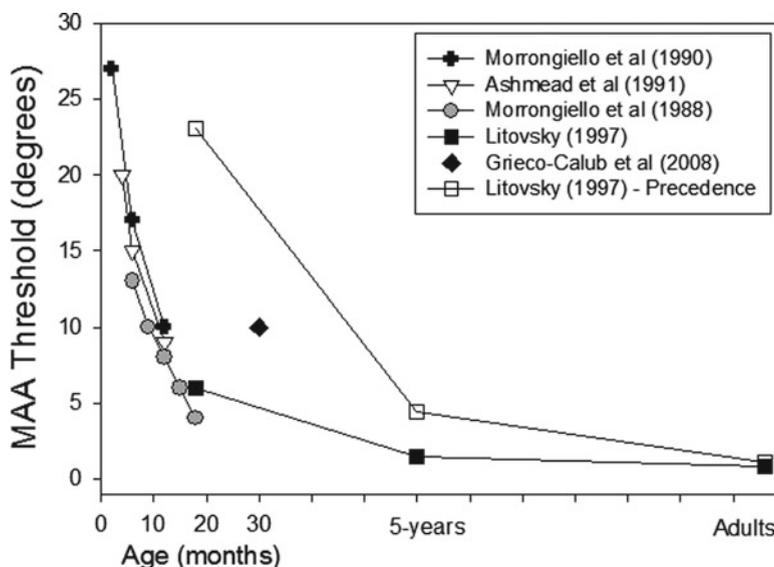


Fig. 6.3 MAA thresholds are plotted in degrees as a function of age in months or years. Data are compared for several studies in which the sound intensity was either held constant from trial to trial (Morrongiello 1988; Ashmead et al. 1991; Litovsky 1997) or in which overall sound pressure level was randomly varied (roved) over an 8-dB range (Grieco-Calub et al. 2008)

to bring visual targets into view and to integrate auditory and visual information. At approximately 3–5 months of age the conditioned head-turn behavior emerges, whereby the infant's response can be shaped such that a reinforcing stimulus is associated with the behavior (Muir et al. 1989).

Figure 6.3 shows the developmental trend in right–left discrimination ability, as measured with head-orienting responses. Data represent MAA thresholds from several studies in which the sound intensity was held constant from trial to trial. From these studies, one can conclude that the largest decrease in MAA occurs between 2 months of age and 2 years of age, with continued improvement through 5 years of age, when children's thresholds are similar to those of adults (see Litovsky 1997).

The diamond data point was collected under a condition in which the overall sound intensity was randomly varied (roved) over an 8-dB range. The consequence of roving level when sound sources are lateralized is that monaural level difference cues become unreliable, and listeners are thus forced to rely more heavily on interaural cues to determine sound source locations. As the data from Grieco-Calub et al. (2008) with 2.5-year-olds show, MAA thresholds are higher when overall level is roved than when the level is held constant (compare, e.g., with MAA in unroved condition from Litovsky's 18-month-olds). When a single reflection is simulated from the front location (see data from Litovsky 1997), MAAs are also higher for infants and children, but not for adults. Thus, spatial hearing acuity develops rapidly

during the first 1–2 years of life in normal-hearing infants, and continues to develop into late childhood under some stimulus conditions.

The MAA is a robust tool for exploring developmental changes in sound localization acuity. The limitation of the MAA task is that it does not capture the extent to which listeners possess a spatial map of the acoustic environment. The fact that a listener is able to discriminate changes in source position does not necessarily imply that the same listener can identify the actual location of sound sources. For example, Wilmington et al. (1994) measured various aspects of binaural processing in children and adults with asymmetric conductive hearing loss. Following corrective surgery, binaural hearing was restored, but performance on tests of binaural processing varied across tasks. These findings suggest that binaural processing is mediated by a complex set of mechanisms, and that although listeners may detect ITD differences, they might not have the proper mechanisms for utilizing the cues in sound localization. These results also have implications for the interpretation of infant MAA threshold; infants appear to be sensitive to ITDs on the order of tens of microseconds (Ashmead et al. 1991), consistent with fairly small MAA thresholds by age 18 months. However, this sensitivity may not translate directly to a spatial hearing map that infants can use to identify source locations or to separate target speech from maskers.

3.2 *Sound Localization Abilities in Children*

Whereas discrimination measures such as the MAA have been utilized with young listeners for several decades, a small number of studies using absolute source location identification methods have been conducted in recent years. Ironically, in some cases, these studies seem to have been motivated by the need for establishing norms and benchmarking development of sound localization accuracy in normal-hearing children so that comparisons can be made with performance of children who are fitted with hearing aids or cochlear implants or both. The motivation for providing hearing in both ears to hearing-impaired children is discussed in more detail later, and in this section some of the sound localization studies in young children are reviewed. There is a range of locations from which to choose (between 7 and 15 in the studies discussed here) and children are expected to be able to either point to the source location, verbally identify the location, or use a computer interactive game to select the location on a screen.

Sound localization studies generally show that children between the ages of 4 and 10 years can localize sounds with average error rates ranging from less than 5° to greater than 30°. Data from three children are shown in Fig. 6.4a (top three panels). For the best-performing child (right panel) with RMS of 9°, mislocalization errors were generally ones in which the response was one loudspeaker location away from the source. The worst-performing child (left panel) had RMS of 29° because, though localization was excellent in the left hemifield (negative location values), there were substantial errors in the right hemifield. A child with intermediate error types for this group is shown in the middle panel with RMS of 12°. In this study (Grieco-Calub

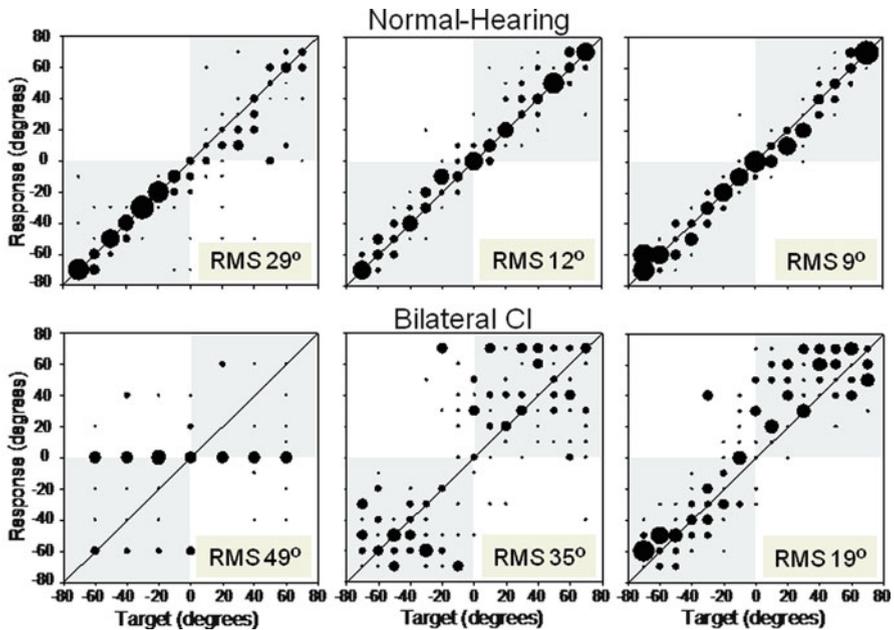


Fig. 6.4 Sound localization results from three children with normal hearing (*top*) and three children with bilateral CIs (*bottom*) are shown. In each panel, the children's responses are plotted as a function of the source locations, and the RMS error values are inserted at the *bottom* of the panel. The size of the *filled circles* is proportional to the number trials on which responses occurred at a particular location. At the *bottom-right* corner of each panel, the child's root-mean-square error is indicated in degrees [From Grieco-Calub and Litovsky (2010)]

and Litovsky 2010), across 7 children age 5 years, root-mean-square (RMS) errors ranged from 9° to 29° (average of $18.3^\circ \pm 6.9^\circ$ SD). RMS errors were generally smaller in two other studies. Litovsky and Godar (2010) reported RMS errors ranging from 1.4° to 38° (average of $10.2^\circ \pm 10.72^\circ$ SD), which overlapped with that observed with adults. Van Deun et al. (2009) found that RMS errors averaged 10°, 6°, and 4° for child ages 4, 5, and 6 years, respectively.

Finally, most recently Litovsky and colleagues have begun to explore the ability of sound localization in 2-year-old children using a nine-loudspeaker array, with sources positioned every 15°. A preliminary report on data from fifteen children (Litovsky, 2011) suggests most children identify locations correctly on greater than 95% of trials (RMS errors $<10^\circ$) while a small number of children reach incorrectly more often (RMS errors near 30°). Of great interest is the fact that some of these children perform in the range of what was observed with the 4–5-year-olds, suggesting that sensory aspects of auditory processing may have matured by this age. However, for some of the children it is likely that factors related to performing on the task itself are still undergoing maturation.

Numerous factors may contribute to the variability in performance seen across studies discussed above on sound localization in children, such as the stimuli, number of loudspeaker locations, angular separations of these locations, and possibly

different response methods or training protocols. Most notable are the facts that (1) Van Deun and colleagues had a much longer stimulus, allowing for head movement cues, and (2) they did not rove the overall sound level. Thus, the children in their study had the advantage of being able to rely on monaural level cues in addition to the binaural cues.

Across these studies it is also clear that children vary in their abilities to perform tasks of absolute localization; whereas some children's performance is well developed by age 4–5, fairly similar to that of adults, other children find the task to be difficult and display large errors in location identification. The extent to which this finding reflects individual differences in sensory as opposed nonsensory factors needs to be explored further. It is unlikely that children would differ in the maturation of the sensory apparatus. Non-sensory factors are more likely to be responsible for individual differences or age-related differences. Although the studies have not been designed to look at this, children are likely to have moments of inattention that add variance to the data. In addition, there may be task-related issues that limit performance in some children; for instance, there may be an issue of mapping perceived locations to the response options on the computer screen.

There is evidence from the animal literature to support this notion. Cats were trained to either look in the direction of a sound or to initiate a behavior such as walking toward the location of the sound source. Ablation of the auditory cortex resulted in an inability to initiate a behavior toward the correct source location, rather than impaired perception of the location (Whitfield et al. 1972). More recently, ferrets have been used as an animal model of the effects of unilateral occlusion on spatial hearing. Most notable are two findings. First lesions of the auditory cortex (A1 in particular) lead to reduction in the ability of animals to relearn spatial hearing maps (Nodal et al. 2010). Second, their behavior improves most dramatically with training and feedback (Kacelnik et al. 2006), suggesting again that nonsensory factors are likely to be involved in the emergence and preservation of spatial hearing maps.

3.3 Sound Localization in Infants Using the Reaching-for-Sound Approach

The first study in which the “reaching in the dark” paradigm was used to study spatial hearing was that of Perris and Clifton (1988), who tested 7-month-old infants using five possible speaker locations in the frontal hemifield. Infants initiated reaching behaviors after the onset of the stimulus on 77% of trials, and their first hand movement consisted of a reach into the correct target area on 75% of the time. Follow-up studies showed that infants also reach for sounding objects in the dark differentially when sound pressure level is varied. Clifton et al. (1991) found that by 6 months of age, infants discriminated between sounding objects presented either within reach or beyond reach when no visual cues were available, suggesting that auditory distance perception may have emerged by that age.

Litovsky and Clifton (1992) subsequently investigated whether infants relied on sound intensity cues. They varied the sound pressure level for near and far objects

so as to eliminate the consistent overall level cue, and showed that infants reached for near objects more frequently than for far objects, regardless of whether the sound pressure level was high or low. This contrasted with adult listeners' verbal discrimination judgments as to whether sounds with low or high intensity were perceived to be far or near, respectively. Thus, whereas adults relied on intensity cues for distance perception, infants did not seem to have acquired a dependence on that cue and were able to use more subtle cues, perhaps spectral or reverberation cues.

The research on "reaching in the dark" suggests that within the first half-year after birth, infants have developed at least a rudimentary spatial map that is functionally meaningful, as they are able to use auditory cues to bring objects of interest into grasp. Further research on this topic is necessary to determine which cues are being utilized by infants at various stages in development, by conducting studies in which cues are eliminated or rendered inconsistent. As discussed previously, infants demonstrate an ability to discriminate small changes in auditory cues that are important for spatial hearing, such as ITDs (Ashmead et al. 1991). However, the ability to discriminate a change in a stimulus feature does not indicate the ability to use that cue for mapping space into a perceptual continuum. That is, ITD discrimination ability or small MAA thresholds do not automatically mean that the infant can use those cues to localize sounds accurately.

3.4 Precedence Effect

As described earlier, when a sound is produced in a reverberant environment it propagates in many directions and is subsequently reflected from nearby surfaces; hence a listener receives both the direct sound and multiple reflections of that sound. One can think of this scenario as involving competition between a sound source and its reflections for perception and localization. Despite this clutter of information, two-eared animals are remarkably accurate at identifying the location and content of sound sources. This reduction in interference from reflections is thought to be related to the PE. Perceptual studies on the PE generally involve a simplified version of the multiple-source scenario that occurs in real rooms. That is, one source (lead) and one echo (lag) are simulated, with a variable interstimulus delay. Unlike the ability to localize single-source stimuli, as described earlier, the PE paradigm offers the opportunity to evaluate conditions under which greater demands are placed on neural coding of timing and dual-source integration. As discussed later, the PE has a more protracted period of maturation than single-source sound localization; thus measuring the PE in children may provide an index for maturation of central auditory function.

Investigations of the PE have used the methods for testing infants and children described in the preceding text. In newborns, unconditioned head turns are elicited. When a single-source sound is presented, within hours after birth they seem able to turn their heads towards the right versus left correctly. However, when lead-lag pairs of stimuli are presented such that the lead is to one side and the lag toward the

other side, it would be the case that a listener who has functional PE would hear the image on the side containing the lead, that is, that the lead would dominate the perceived location, a phenomenon known as localization dominance (see Sect. 1.2.4). Localization dominance has not been found at birth (Clifton et al. 1981), that is, infants do not orient their heads toward the location of the leading source. This has been interpreted to indicate that they are unable to weight more heavily directional cues related to the leading sound. This finding suggests that lagging sounds or “unheard echoes” can exert a strong influence on directional hearing in young listeners.

Localization dominance, measured with the conditioned head turn response, begins to emerge at age 4–5 months (Muir et al. 1989). However, the manner in which the brain handles echoes continues to develop into middle-to-late childhood. For instance, infants do not seem to respond to the echo as a separate sound until the delay is increased to more than 25 ms, which is significantly higher than the values reported for adults under the same conditions. This finding might lead one to conclude that the PE is “beneficial” to the child early in life because the echo is unheard and should therefore not interfere with directional hearing. However, data on directional hearing in the presence of unheard echoes suggest that the opposite is true. Although infants may not show a robust response to the lagging sound, the lag can nonetheless be quite disruptive to localization of the leading sound, perhaps by exerting a “pulling effect” on the perceived location of the lead. Using MAA to measure localization acuity in young children, Litovsky (1997) found that the presence of a lagging sound at delays below echo threshold (i.e., when it is not heard as a separate sound) degraded localization acuity significantly. This can be seen in Fig. 6.2, in which the data from Litovsky (1997) in the precedence effect condition are plotted along those from single-source measurements. MAA thresholds in the precedence effect condition are much greater than single-source MAAs, for the two groups of children tested (18 months and 5 years), but not for the adult group. Thus, there appears to be an immature ability in young children to suppress directional information from a delayed sound that is not perceived as a separate auditory event. In the absence of the delayed (unheard) stimulus, MAA acuity appears to be mature by age 5. However, results from conditions in which the delayed stimulus was used highlight the fact that spatial hearing acuity in children is vulnerable to degradation by virtue of the fact that the auditory system is unable to dismiss directional cues carried by the delayed stimulus.

The simple stimulus scenario used in studies of the precedence effect is by no means representative of the stimuli that one might encounter in everyday listening situations, because here only a single reflection is simulated and its intensity is the same as that of the original sound source. Nonetheless, the fact that a single delayed sound can disrupt localization acuity suggests that careful consideration should be given to children’s abilities to extract information from stimuli in more realistic, complex and reverberant acoustic spaces.

More recently Litovsky and Godar (2010) studied localization dominance (the extent to which the leading source dominates sound localization) and its relation to fusion (the lead-lag delay at which listeners perceive the lagging source as a

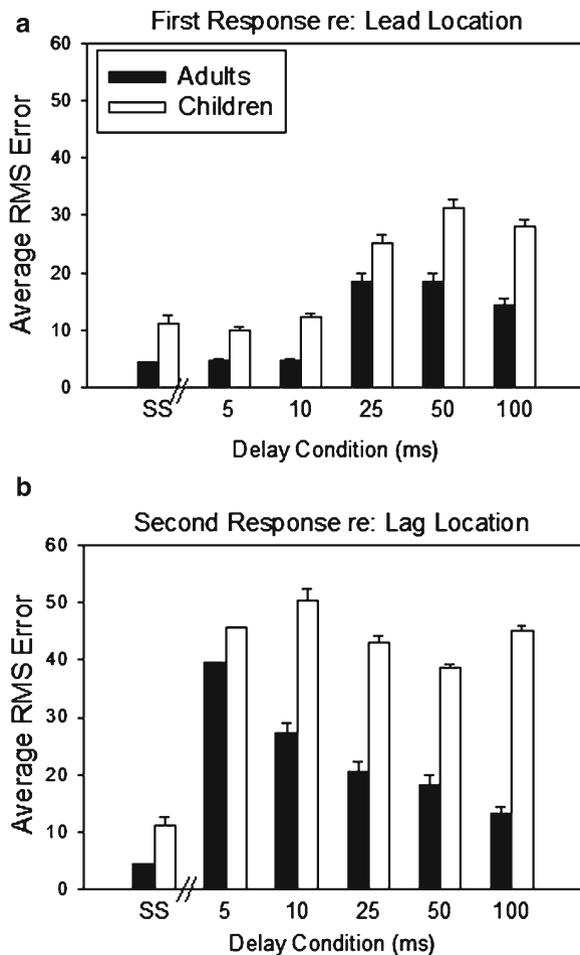


Fig. 6.5 Results from a precedence effect experiment, in which the delay for lead-lag stimulus pairs varied from 5 to 100 ms, and results are compared with the control single-source (SS) condition. Lead and lag locations were as shown in Fig. 6.2b. In each panel, results are compared for groups of adults and children. (a) RMS error indicates the accuracy with which listeners were able to localize the lead source relative to its true location. (b) RMS error indicates the accuracy with which listeners were able to localize the lag source relative to its true location at 0° [From Litovsky and Godar (2010)]

separate auditory event). A unique experimental paradigm was used, whereby listeners indicated how many sources they heard, and reported the perceived location of each source. This study captured the ability of listeners to localize the leading sound and the lagging sound at delays that were below fusion echo threshold and perceived as a single fused auditory event. In addition, the study investigated the ability of listeners to perceive the lead and lag at longer delays, above echo threshold, when the sounds are generally perceived as being separate. Stimulus configurations

for these stimuli are shown in Fig. 6.2, and the locations of the lead and lag are shown in Fig. 6.2b. Children ages 4–5 years and adults participated. Results from the fusion measure, in which listeners reported whether they heard a single (fused) sound or two sounds, suggested that children had significantly higher fusion echo thresholds (24.3 ± 13.6 ms) than adults (15.2 ± 4.9 ms). Figure 6.5 shows localization data with RMS errors for localizing the lead, or lag. RMS errors for lead localization (top panel) were fairly similar to errors observed with single-source stimuli, when the lead-lag delays is short, suggesting that the presence of an “unheard echo” does result in worse sound localization. As delays increased, RMS errors for the lead increased. That is, lead localization worsened, for both adults and children, suggesting that a “heard echo” disrupts localization.

In addition, RMS errors were consistently higher in children than in adults, suggesting that their ability to suppress or ignore directional cues from the lagging source was not as fully developed as that of adults. The implication of this finding is that in reverberant environments children may have more difficulty than adults in negotiating spatial cues arriving from sources and their reflections. The ability to localize the lag (bottom panel) also showed interesting age-related differences. Both adults and children had large errors at short delays (i.e., when fusion was strong). At longer delays, adults’ abilities to identify the lag location improved with increasing delay, whereas children’s error rates remained high at all delays. This finding suggests that children had a general inability to resolve locations of two identical stimuli whose only difference is onset time. The authors interpreted the findings within the context of temporal-order confusions, localization interference, and over-integration of the two images despite their asynchronous onsets.

What factors could account for immature PE in children and infants? Sensitivity to simple (i.e., single-source) stimulus configurations can be modeled by invoking fairly low-level processing in the binaural pathway (Xia et al. 2010), whereas localization of PE stimuli requires higher-level (central) processes. Neural inputs that are initially involved in establishing the PE probably exist at peripheral (Hartung and Trahiotis 2001) and brainstem (Litovsky and Yin 1998; Litovsky and Delgutte 2002) levels. However, more central (cortical) mechanisms seem to be necessary for mediating behavioral components associated with the PE (Clifton et al. 1994). Animal ablation-behavioral work has shown that unilateral ablation of the auditory cortex severely impairs the ability of cats to accurately choose the hemifield containing the leading source, while the ability to find the hemifield containing a single-source sound was not affected (Cranford et al. 1971; Whitfield et al. 1972). Performance was consistent with elimination of lead dominance in localization on the side of the lesion. The deficit was limited to cases in which the leading source was in the contralateral hemifield (opposite to the lesion site). In addition, whereas binaural circuits in the newborn cat show neural correlates of the PE that are similar to those seen in the adult brain (Litovsky 1998), the behavioral immaturity of the PE in infants (and other species; e.g., Ashmead et al. 1986) is not be easily explained by processing at the level of the brainstem. Developmental differences observed on the PE task may share underlying mechanisms with other perceptual phenomena that require binaural processing, such as temporal integration (see Sect. 3.5).

3.5 *Binaural Sensitivity and Binaural Unmasking*

The developmental time course of perceptual abilities described in this chapter thus far reflects measures that have been conducted in the sound field, when listeners receive a multitude of acoustic cues at the ear drums. Thus, a combination of monaural (level and spectral) and binaural cues is available to listeners for functional purposes. A different approach can be taken to investigating aspects of binaural hearing that develop during childhood, by focusing on sensitivity to specific sets of cues. That is, by presenting sound through headphones, one can manipulate the acoustic cues reaching the ear drums such that some, or all, of the cues from the sound field are represented in the acoustic signal. As such, binaural sensitivity can be studied under conditions in which the difference between the ears in timing and level of the signals are controlled and varied systematically. Developmental studies on binaural sensitivity are sparse. Ashmead et al. (1991) reported that by 4–7 months of age human infants are sensitive to ITDs on the order of 50–75 μ s. Although some what higher than the best ITD threshold of less than 10 μ s obtained with highly trained adult listeners, this range of performance is similar to that seen in untrained adults (Wright and Fitzgerald 2001). A recent study by Van Deun et al. (2009) measured ITD sensitivity in children ages 4 years and older, who are perhaps more trainable than infants, and reported ITD thresholds of: 40 μ s for 4-year-olds, 20–35 for 5–9-year-olds, and 12.5 μ s for adults. Thus, by age 4 years or perhaps slightly older, ITD sensitivity is fairly well established. This is consistent with findings discussed earlier regarding children’s abilities to localize sounds in the free field and to discriminate source locations in a right versus left task (MAA), where data suggest that performance is generally close to that seen in of adults.

As discussed in Sect. 1.2.3, BMLDs are used as a measure of binaural sensitivity, with some attempts to relate these measures to speech unmasking. Numerous studies with adults have shown that a signal can be heard in noise in the NoS π condition (different binaural cues to the signal and noise) because of dissimilarity or decorrelation between the signal and noise at the ears. Nozza et al. (1988) reported BMLDs for speech stimuli of 5, 8, and 10 dB for infants, 4-year-olds, and adults, respectively, and attributed the developmental effects to age differences in absolute sensitivity in masked conditions, rather than reduced ability of younger listeners to segregate target from noise. Studies in which the signal was a 500-Hz tone masked by noise suggest that children between the ages of 4 and 12 years have BMLDs similar to those of adults for wideband noise burst maskers, but smaller BMLDs for narrowband noise maskers. In addition, younger children (~4 years) are generally slightly worse than older children or adults (~7–12 dB depending on the noise bandwidth; Grose et al. 1997; Van Deun et al. 2009).

The BMLD paradigm has also been utilized to investigate developmental differences in “binaural sluggishness” (e.g., Grantham and Wightman 1978; Kollmeier and Gilkey 1990). Consider the time interval over which the auditory system integrates information related to binaural difference cues. In children ages 5–10 years, the integration occurs over a longer time window than in adults, on the order of

40 ms after onset of the signal (Hall et al. 2007). In fact, this temporal window is similar to the lead-lag time delay over which localization confusions occur with the PE stimuli (Litovsky and Godar 2010). Adults may be more adept than children at utilizing smaller and more optimal temporal windows during which they can listen and optimize the ability to extract information from relevant stimuli. Further, children may be less able to focus their listening efforts over a small time interval, which effectively leads to integration of information from stimuli that fall into these time intervals (windows).

Taken together, these studies suggest that children are able to utilize binaural cues to segregate target tones from maskers, and the effects can be rather large. However, there are factors that reduce the ability of children to gain the same effects as adults. These factors are nonsensory, perhaps cognitive, or related to listening strategy, as well as ones that depend on experience.

4 Psychophysical Studies on the Speech Intelligibility in the Presence of Maskers

Given that children are clearly able to utilize binaural cues to segregate signal from noise, it would be reasonable to expect them to also be able to segregate speech from maskers in realistic listening situations. However, in that case they must be able to make sense of the “what” and “where” aspects of the auditory environment. The phenomenon by which the auditory system extracts a distinct message in the presence of other competing sounds is known in general terms as the “cocktail party effect” (Cherry 1953; Pollack and Pickett 1958), aspects of which have been simulated and studied extensively in adults (see discussion earlier in this chapter). Although there has been increasing interest in identifying the mechanisms by which the auditory system of adult listeners teases apart co-occurring sounds and facilitates speech understanding in noisy environments, our understanding of the developmental aspects of these effects is in its infancy. Perhaps the issue that renders these studies most challenging is the fact that, in conducting developmental research, we can best contribute to scientific progress if we can tease apart sensory processes from nonsensory ones, such as attention, cognition, and central auditory development. This issue has been discussed earlier with regard to age-related differences and within-age variability in the performance of children on spatial hearing tasks.

A number of studies by Litovsky and colleagues have simulated aspects of the auditory environment that might be encountered in a realistic listening situation, fondly dubbed the “classroom party effect.” Litovsky (2005) first demonstrated spatial release from masking (SRM) in children ages 4–7. Target stimuli were presented from the front at 0°. Maskers were presented from locations that were either collocated with the target, or spatially separated from the target. The main difference between adult and child studies is that, for children, a four-alternative forced-choice (4AFC) task is implemented in order to enable children to indicate what they hear rapidly and reliably. Before testing, children are familiarized with target spondaic words, selected

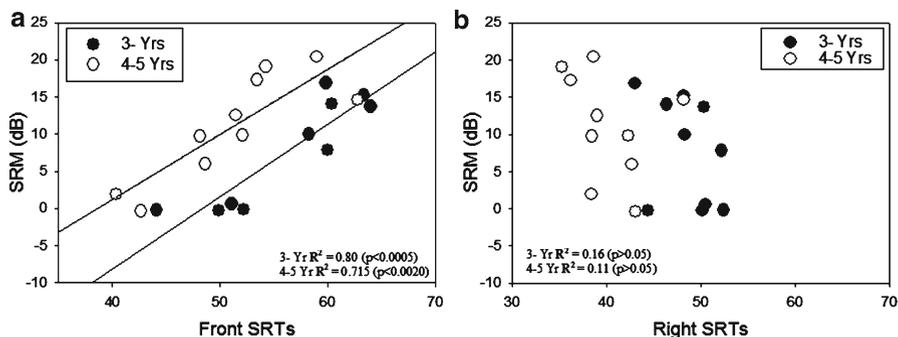


Fig. 6.6 Values of spatial release from masking (SRM) for individual children are plotted as a function of the speech reception threshold (SRT) either when the masker was in front, colocated with the target (a), or when the masker was on the *right*, spatially separated from the target (b). Within each panel, data from 3-year-olds and 4–5-year-olds are compared

such that they could each be represented with a visual icon (e.g., ice-cream, cow-boy, bird-nest), and these speech stimuli are within the vocabulary of typical 4-year-old children. Maskers consisted of sentences strung together that do not overlap in content with the target speech. This work is thus different from other speech unmasking paradigms in which the target and maskers are constructed from the same materials and easily confusable, resulting in informational masking (e.g., Brungart and Simpson 2002; Wightman and Kistler 2005).

Litovsky (2005) found that in the colocated condition, children experienced masking that averages 16 dB and 21 dB in the one- and two-masker conditions, respectively. When the masker was displaced to 90°, children experienced spatial release from masking (SRM) averaging 5.2 dB and 7.4 dB, in the one- and two-masker conditions, respectively. Thus, children are able to benefit from differences in spatial cues between target speech and masker. It is worth noting that the target-masker configurations that result in SRM were due to a combination of acoustic “head shadow” cues and the binaural cues, interaural time, and level differences. This combination of cues occurs when the masker is on one side of the head, and the target-to-masker level ratio favors the target in the ear that is farther from the maskers. Misurelli et al. (2010) recently reported that children ages 4–7 years demonstrate SRM even in absence of the head shadow cues, that is, when the maskers were displaced toward both the right and left, thus resulting in a condition with no “better ear.” However, the SRM was smaller than when both head shadow and binaural cues were present (Misurelli et al. 2010).

Garadat and Litovsky (2007) examined SRM in 3-year-old children. Target words were chosen to be within the receptive language and vocabulary of children ages 2.5–3.0 years. As in Litovsky (2005), the child selected a visual icon to match the heard word on each trial. By age 3 children had SRM values that were similar to those of 4–5-year-olds, suggesting that the ability to benefit from spatial separation between target and maskers developed at this young age (Fig. 6.6). Further, children

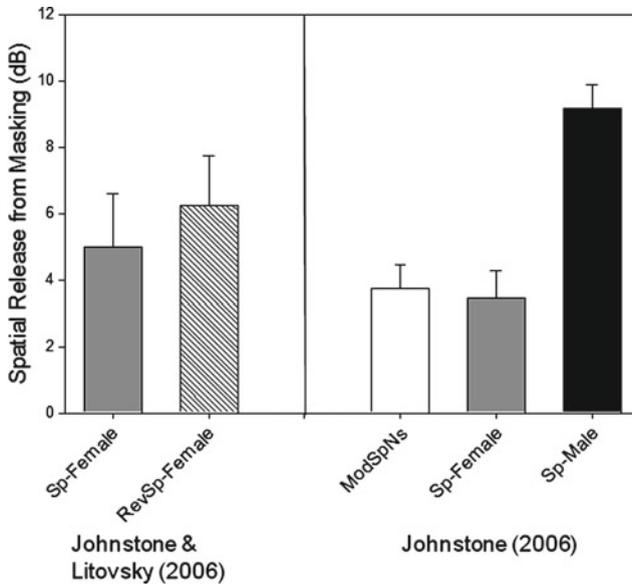


Fig. 6.7 Spatial release from masking (SRM) values are shown for two studies in which 6–7-year-old children were tested. In each panel, SRM is plotted in decibels, as the difference in speech reception thresholds (SRTs) between conditions in which the target-masker are collocated versus spatially separated. Positive SRM values indicate that SRTs were higher (i.e., performance worse) in the collocated than the separated conditions. *Left panel* (data from Johnstone and Litovsky 2006): SRM is compared for conditions in which the masker was either speech spoken by a female talker (Sp-Female, gray) or time-reversed speech constructed from the female speech (RevSp-Female, hatched). *Right panel* (Data from Johnstone 2006): SRM is compared for conditions in which the masker was either modulated speech-shaped noise (white), speech spoken by a female talker (Sp-Female, gray) or speech spoken by a male talker (Sp-Male, black)

who demonstrated the greatest SRM were those with high speech reception thresholds (SRTs) in the Front (0°) condition where the target and maskers were collocated. Figure 6.6a shows the correlation between SRTs in front and SRM, and Fig. 6.6b shows that SRTs with masker on the side do not correlate with SRM.

In the Litovsky (2005) study, SRM was shown to be larger when two-masker speech stimuli were used than when one-masker was used, suggesting that, when the listening environment is more complex, spatial cues play an increasingly more important role in assisting the listener with source segregation. Several follow-up studies have been conducted in which other aspects of the auditory environment have been manipulated, such as the type of masker. Figure 6.7 shows SRM values for four different types of maskers from two studies with children ages 6–7 years (Johnstone 2006; Johnstone and Litovsky 2006). In Johnstone’s (2006) study, SRM was greatest when target-masker similarity was high (i.e., when both the target and masker were produced by a male talker). Studies with adult listeners have reported similar findings (Bronkhorst 2000; Brungart and Simpson 2002). This effect has been attributed to the fact that target-masker similarity renders speech segregation

extremely difficult, making spatial cues the most salient cues that listeners can use to segregate target from maskers. By comparison, SRM is smaller with different-sex and nonspeech maskers, because cues resulting from differences in the spectra of the signals can be useful for source segregation.

A somewhat different variation on the target-masker similarity can be introduced with a reversed-speech masker that contains the same temporal amplitude fluctuations and long-term spectrum as speech. In children, this masker produced SRM that was similar to that with the speech masker from which it was created, even though it carried no linguistic content (Johnstone and Litovsky 2006), a finding that is consistent with observations in adults (e.g., Hawley et al. 2004). Children provided anecdotal reports that the masker had a novel feature, with some resemblance to a person speaking in a foreign language. This may have added to the interference that would have been produced simply by a modulated speech-shaped noise masker that bears similarity in the spectral and amplitude-modulation domain but carries no resemblance to spoken language.

Another interesting issue is whether the side (right vs. left) to which maskers are displaced has an impact on SRM. In general, maskers have been placed on the right in spatially separated conditions. Johnstone (2006) studied the effect of masker side; SRM in adult listeners was smaller for conditions with maskers on the left, that is, when the right ear had a better target/masker ratio. The asymmetry was particularly robust when trial-by-trial uncertainty was introduced by varying the content of the masker (speech, reversed speech, modulated noise). These data are consistent with a recent imaging study in adults showing that on speech-listening tasks involving informational masking there are unique activation patterns in the right-brain auditory fields (Scott et al. 2009). In contrast with adults, the 30 children ages 6–7 years tested by Johnstone did not show the spatial asymmetry. A glimpse into developmental trends was observed when the data were separated by gender: Girls showed a small asymmetry in the same direction as the adults, an effect that was obscured by the mixed-gender analysis. The issue of gender differences in complex listening conditions has not been explored and may be an important avenue of research to pursue.

Effects of talker gender and developmental trends in speech masking have been found under conditions of informational masking (Wightman and Kistler 2005; see also Leibold, Chap. 5). In this paradigm, stimuli were presented over headphone and the ability to selectively listen to one ear and ignore maskers that either in the same ear or in the opposite ear was measured. Informational masking was induced by using target and maskers from the same stimulus corpus, that is, with a high chance for target/masker confusability. When target and maskers were presented monaurally to the same ear, children demonstrated greater informational masking than adults. This age difference in masking was not observed by Litovsky (2005), most likely because that study had different speech materials and different-sex talkers for target and masker, hence little or no informational masking. Wightman and Kistler (2005) also found that adding a distracter message in the opposite ear resulted in reduced performance that was much greater for children than adults, and that although performance improved with age, even the oldest age group (up to 16 years) did not perform similarly to adults.

5 Binaural and Spatial Hearing in Special Populations

A growing number of children with hearing impairment have been receiving stimulation in both ears in an effort to provide them with perceptual benefits on auditory tasks that are known to rely on having inputs in both ears. Bilateral hearing aids have been the standard of care for amplification for many years (Litovsky and Madell 2009). Another population of children has been receiving cochlear implants (CIs), which are successful in providing auditory input by electrically stimulating the auditory nerve. For many users this input is sufficient for attainment of speech perception and production that are within normal range, allowing aural–verbal communication. Adequate speech perception and language acquisition can be achieved through the use of a single implant. However, many parents of children who are eligible for CIs are electing for their child to be implanted bilaterally to improve the child’s ability to segregate speech from background noise and to localize sounds.

The vast majority of adults who became deaf postlingually and are implanted bilaterally as adults show significantly better sound localization abilities when using both CIs than when using a single CI (van Hoesel and Tyler 2003; Litovsky et al. 2009; van Hoesel 2004). Typically, these adults will have had exposure to acoustic hearing for many years before becoming deaf, and the activation of bilateral CIs most likely reactivates some aspects of their previously established spatial-hearing abilities. In children the issues are quite different, because most of them will not have been exposed to acoustic input before becoming deaf. Rather, they are typically born deaf and prior to implantation continue to undergo a period of deafness during which there is little if any, auditory input. Subsequent to becoming implanted they undergo a period of adaptation to electrical stimulation. Studies with this population offer an opportunity to examine effects of auditory deprivation and unusual (electric) auditory inputs on the development of spatial and binaural hearing.

When considering spatial hearing abilities in children who are deaf and use CIs, it is important to do so in the context of what we have learned about these same abilities in children who have normal hearing. One is faced, however, with an interesting issue regarding controlled comparison groups. One can compare performance between children who are normal hearing and who use CIs, at the same chronological age, but the normal-hearing children will have had more “time in sound” because of the early period of deprivation that CI users experience. Another option is to compare children who have had the same amount of “time in sound,” in which case the CI users will be chronologically older. In reality, neither of these approaches has been rigorously implemented because of limitations in the subject population under consideration and the small number of studies conducted to date.

Litovsky and colleagues have been studying the emergence of spatial hearing skills in young children who are fitted with bilateral CIs. The vast majority of the children receive their two CIs in sequential procedures, separated by a period that can last between several months to several years. Behavioral measures in these children have generally been similar to those obtained in children with normal hearing. The goals of the research have been to (1) evaluate whether bilateral CIs provide a

benefit over a single CI and (2) determine the extent to which performance in the CI users is similar to that observed in normal-hearing peers.

A caveat that will appear several times in the text that follows is the inherent limitation in the speech processors that are used in today's CIs. The hardware and signal processing in the implantable devices are far from ideal as far as providing binaural cues with fidelity. Bilateral CI users are essentially provided with two separate monaural systems. Speech processing strategies in clinical processors utilize pulsatile, nonsimultaneous multichannel stimulation, whereby a bank of bandpass filters is used to filter the incoming signal into numerous frequency bands (ranging in number from 12 to 22), and to send specific frequency ranges to individual electrodes. The envelope of the signal is extracted from the output of each band and is used to set stimulation levels for each frequency band. Thus, fine structure is discarded. Although ITDs in the envelopes may be present, because the processors have independent switch-on times, the ITD can vary dynamically and unreliably (van Hoesel 2004). In addition, the microphones are not placed in the ear in a manner that maximizes the capture of directional cues such as spectrum and level cues. Microphone characteristics, independent automatic gain control, and compression settings distort the monaural and interaural level directional cues that would otherwise be present in the horizontal plane. Binaural speech processors do not provide bilateral CI users with similar cues available to normal-hearing listeners. The eventual manufacturing of such clinical processors will presumably take into account the extent to which CI users are sensitive to, and able to utilize binaural cues.

5.1 Spatial Hearing Acuity in Children Who Use Bilateral CIs

As reviewed earlier, in normal-hearing children rudimentary spatial hearing abilities are observed at birth, with the unconditioned head turn response. During the first few years of life, spatial hearing acuity emerges and improves such that MAA thresholds are only a few degrees by 2 years of age, and similar to adults by age 5 years (see Fig. 6.2). Using similar approaches, Litovsky et al. (2006b) studied children ages 3.5–16 years who were bilaterally implanted; that study provided the first evidence suggesting that bilateral CIs can result in better spatial hearing acuity than unilateral CIs. MAA thresholds were measured when the children were listening through both of their CIs, or when only one CI was activated. Having access to stimulation in both ears resulted in smaller MAA thresholds compared with the single-CI condition. Nine of thirteen children tested were able to perform the MAA task above chance, most of them achieving MAA thresholds of less than 20°. Although thresholds were overall higher than those observed in age-matched normal-hearing peers, the goal of the study was to compare performance with two CIs versus one CI; eight of nine children performed better when using both CIs than when using a single CI. Interestingly, the remaining 4 of 13 children were not able to perform the task and (anecdotally) appeared to have little understanding of the fact that sounds can carry information regarding spatial location. It is likely that degraded

spatial hearing ability was affected by lack of binaural input at a time when spatial hearing sensitivity would have developed. In addition, even after the children were activated bilaterally they were unlikely to have been provided with binaural cues that are reliable, because of inherent limitations in the CI speech processor. This issue is discussed in more detail in the preceding text.

In a recent study with a similar group of sequentially implanted children, Godar and Litovsky (2010) were interested in the change that occurs in MAA when children transition from using a single CI to using bilateral CIs. Children were tested before activation of the second CI, and again at 3 months and 12 months following bilateral activation. Results from this study are shown in Fig. 6.8a. MAA thresholds generally improved from preactivation to the bilateral test at 3 and 12 s. It is important to note that unilateral performance did not improve, as expected, in part because unilateral spatial cues are not too useful on spatial hearing tasks such as this one. It may also be the case that unilateral performance did not improve because these children did not continue to use unilateral CIs in their everyday listening environment, and any learning involving use of input from the CI that might occur on a daily basis would occur only in the bilateral listening mode.

The aforementioned studies on emergence of spatial hearing acuity in children who are deaf and used bilateral CIs found that, in general, implanted children's performance asymptotes at MAA thresholds of 10–20°, which is considerably higher than the values of 2–5° that have been measured in normal-hearing age-matched peers. One of the factors that differentiate between the implanted and normal-hearing children is the fact that implants are typically provided during the first few years of life, and the young recipients will have spent months or years with no auditory input. This factor led Grieco-Calub et al. (2008) to investigate whether younger implanted children, who will have undergone shorter periods of auditory deprivation before activation of bilateral hearing, will “catch up” with their age-matched peers. Grieco-Calub et al. (2008) tested spatial hearing acuity in toddlers, who had received their first CI by age 12 months and second CI by age 18 months significantly earlier activation of bilateral hearing than in previous studies. In addition to testing bilaterally implanted children under conditions in which either one or two CIs were activated, Grieco-Calub et al. tested a control group of unilaterally implanted children who were accustomed to using one CI in their everyday lives. Finally, a group of eight normal-hearing, typically developing toddlers also participated. Results from that study are shown in Fig. 6.8b. When compared with their normal-hearing peers, the group of children with bilateral CIs showed larger intersubject variability in performance, which is a common feature among individuals who use CIs. Note that three toddlers in the bilateral CI group had MAA thresholds that are within the range observed in their age-matched peers. These findings suggest that early activation of the auditory system in both ears of children who are deaf may result in emergence of spatial hearing abilities in at least some of the children. However, why are some of the children unable to perform the task? One of the most obvious limitations in today's bilateral CIs is the lack of coordination between the processors in the two ears. The absence of consistent and accurate binaural stimulation may account for the inability of some

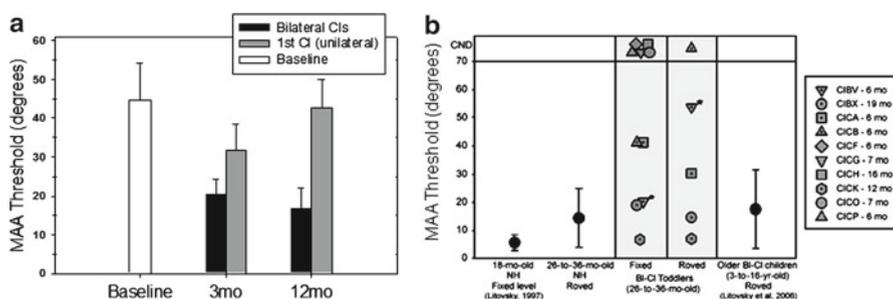


Fig. 6.8 (a) Group mean MAA thresholds for a group of ten children who were tested at three intervals. Baseline is the testing interval during which children either wore a single CI or a CI and a hearing aid in the other ear, before activation of the second CI. Testing was repeated for each child at 3 months and 12 months following bilateral activation. Conditions included bilateral listening mode (*black fill*) and unilateral listening mode (*gray fill*) (From Godar and Litovsky 2010). (b) MAA thresholds for four groups of children. *Left-most*, children with normal hearing age 18 months (From Litovsky 1997; see also Fig. 6.2), with overall sound level fixed. Next, children with normal hearing age 26–36 months, tested with overall sound level roved. Shaded areas in the *middle panels* include data from deaf toddlers with bilateral CIs who completed the task with overall sound level fixed, and a subset of these children who were retested with overall sound level roving (From Grieco-Calub et al. 2008). *Right-most*, average MAA data are re-plotted from a group of older children with bilateral CIs [Litovsky et al. 2006b; see text for discussion]

of the children to learn to localize sounds. Alternatively, it is possible that some of the cues that are useful for sound localization are in fact being presented to these children, but the period of auditory deprivation that they experienced before being activated has the effect of limiting their ability to form a spatial hearing map.

5.2 Sound Localization in Children Who Use Bilateral CIs

Although results from studies on left–right discrimination are meaningful regarding the emergence of spatial hearing, a perhaps more ecologically relevant approach is one in which sound location identification, or absolute localization, is investigated. Grieco-Calub and Litovsky (2010) recently tested a group of 21 children ages 5–14 years who use bilateral CIs. As discussed earlier, studies on sound localization in normal-hearing children between the ages of 4 and 10 years show average error rates ranging from less than 5° to greater than 30° (see Fig. 6.4). In the bilaterally implanted group, 11 of 21 children had RMS errors that were smaller when they used both CIs than when a single CI was activated, suggesting that they experienced a bilateral benefit. Focusing on their bilateral-condition responses, sound localization was highly variable among the 21 children, with RMS errors ranging from 19° to 56° , compared to the 5-year-old normal-hearing group, who had RMS errors ranging from 9° to 29° . Children with bilateral CIs fell into one of three groups, and examples from each group are shown in Fig. 6.4b: children who could not localize (left panel

example; group of $n=7$), children who could discriminate between sources arriving from the left versus right hemifield, but with no clear indication that they can discriminate between source locations within a given hemifield (middle panel example; group of $n=8$), and children who localize with RMS errors within the range of performance observed in normal-hearing children (right panel example; group of $n=6$). In this latter group, although the children's performance fell within the range of performance seen in normal-hearing children, their localization abilities were still worse. The scatterplots suggest that some children have bias toward one hemifield or another and there is considerable scatter around the means. Compared with the normal-hearing children, children with bilateral CIs have many fewer trials with absolute correct identifications. The data suggest that spatial hearing resolution, or internal "map" of space, is perhaps more blurred and less acutely developed in the CI users than in children with normal hearing. What is somewhat remarkable is that the bilaterally implanted children were able to localize sounds at all, given that the hardware and signal processing in the implantable devices are far from ideal as far as providing binaural cues with fidelity.

5.3 Lessons Learned from Adults Who Undergo Prolonged Periods of Auditory Deprivation Before Receiving Bilateral Stimulation

In recent years, there has been growing interest in binaural abilities in bilaterally implanted listeners. As far as auditory development is concerned, one of the most interesting facets of studying people who receive CIs is the opportunity to explore the effect of short- and long-term auditory deprivation on perception, in this case on binaural sensitivity. Several studies have shown that adults whose onset of deafness occurs during adulthood, after having had years of exposure to acoustic cues, are able to retain sensitivity to ITD cues (see van Hoesel 2004 for review). Litovsky et al. (2010) recently studied binaural sensitivity in three groups of adult listeners who had become deaf at various stages in life: (1) adults who became deaf when they were very young children (prelingual); (2) adults who became deaf during their childhood after experiencing acoustic hearing for some years (childhood); and (3) adults whose onset of deafness occurred after they reached adult age (adult). In these experiments, binaural stimulation was carefully controlled: Electrically pulsed signals were transmitted directly to specific electrodes in the electrode arrays of the CIs. Electrodes in the right and left ears were selected so that sounds delivered to those electrodes were matched in pitch and loudness, and so that simultaneous activation of the two electrodes resulted in perceptually fused, coherent auditory images perceived to be in the center of the head. These prerequisites ensured that the binaural system received binaural cues with fidelity. Results suggest that ITD sensitivity is impacted by early deprivation: None of the subjects in the prelingual group had access to ITDs, whereas subjects in the other groups retained sensitivity to ITDs. The primary difference between the prelingual and childhood-onset group was that the latter

heard sound for the first 8–10 years of life, whereas the former recall having had severe-to-profound hearing loss their entire lives. It is difficult to assess how many years of exposure to acoustic input is necessary for the establishment of ITD-dependent binaural sensitivity. It is noteworthy that ILD cue sensitivity was present in all subjects. That is, even the prelingually deafened adults were able to use ILDs to discriminate between sources that were presented to the right versus left, and to perceive laterally displaced images “in the head” when ILDs were presented.

These findings suggest that the mechanisms involved in processing ITD cues are more susceptible to hearing loss than the mechanisms associated with ILDs. From a practical point of view, the lack of sensitivity to ITD may not be critical to functionality with today’s CI processors. This is because most speech processing strategies do not preserve the fine-structure cues that can give rise to ITDs. Rather, CI processors discard fine-structure information and use amplitude-modulated fixed-rate pulsatile stimulation. The resulting stimuli bear some resemblance to acoustic stimuli comprising high-frequency carriers that are amplitude modulated. Because humans are known to be sensitive to ITDs in the envelopes of these high-frequency modulated stimuli (e.g., Bernstein and Trahiotis 2002), there is the possibility that, in bilateral CIs, ITDs in the envelopes can occur for sounds that reach the two ears with difference in arrival time. It is worth investigating whether the effects of early auditory deprivation on binaural sensitivity are limited to ITD processing that depends on low-frequency stimuli and encoding of fine structure. In that case, bilateral CI users with early deprivation may ultimately be able to recover binaural sensitivity by relying on the auditory circuits that encode high-frequency information.

The differential effects of deprivation on ITD and ILD sensitivity may be rooted in the mechanisms by which ITD and ILD are coded and processed in the auditory system. Converging evidence from anatomical, physiological, and psychological studies has suggested that ITD is dependent on anatomical circuitry comprising a neural “display” in which neurons are arranged topographically in the brain stem by the frequency to which they are sensitive. Delays in arrival at the ears of external ongoing stimuli are compensated by internal anatomical wiring, by varying axonal length (Joris and Yin 2007). This specialized processing scheme is observed only in the ITD circuits. In contrast, several brain stem nuclei respond to level differences and there are a number of ways in which ILD processing can be achieved that does not depend on precise timing or axonal length (e.g., Green and Sanes 2005). Finally, synaptic transmission processes appear to be vulnerable to disruption induced by hearing loss in such a way that ITD coding may be more vulnerable to hearing loss than ILD coding (Svirskis et al. 2002).

5.4 Speech Intelligibility in Noise in Children Who Use Bilateral CIs

The extent to which children who are deaf and use bilateral CIs can hear speech in noise has been a topic of great interest in recent years. Because children with normal hearing are able to use information presented to both ears to segregate speech from

maskers in complex listening conditions, on a clinical level there may be an expectation that bilaterally implanted children might perform similar to their peers on these tasks. However, an important consideration is that acoustic information that is available to children who use amplification and implantable devices is unlikely to be “normal,” and the performance of these children is therefore likely to be worse than that of their normal-hearing peers.

One obvious issue is that the two ears of deaf children are often implanted sequentially, and hearing in the two ears is likely to be asymmetrical. The population of children with bilateral CIs is quite varied. Some of the most interesting factors are the age at which bilateral hearing is activated and interactivation time intervals, because the auditory system of each child undergoes periods of auditory deprivation that is unique to the child. In some instances, children are profoundly deaf at birth and implanted simultaneously in both ears, in which case there is little or no difference in the auditory experience of the two ears. More typically, auditory experience is unequal between the two ears. As a whole, these children and their unique experiences offer interesting opportunities to consider the impact of auditory plasticity and maturation on performance.

Peters et al. (2007) compared speech intelligibility outcomes in children who received their second implant by age 3.0–5.0, 5.1–8.0, or 8.1–13.0 years. Remarkably, children were able to obtain open-set speech discrimination in their second-implanted ear even when receiving their second implant as late as 13 years of age. However, performance with the second-implanted ear did not “catch up” to that with the first-implanted ear by 12 months after activation of the second CI in children older than 8 years. In contrast, in children younger than 8 years, speech scores in the second-implanted ear rose more quickly with experience and were nearly as good as first-ear scores within 12 months of bilateral activation. Finally, in general, these children showed better performance when listening bilaterally, even following years of monaural deprivation in the second-implanted ear.

Speech intelligibility with speech maskers similar to those used with normal-hearing children (see Sect. 4) has been used as a measure of bilateral benefit. Two populations of children have been studied: those who use bilateral CIs or those who use a CI in one ear and hearing aid in the other ear (e.g., Ching et al. 2005, 2006; Litovsky et al. 2006a). Average group data from Litovsky et al. (2006a) are shown in Fig. 6.9. In general, results suggest that having access to inputs in both ears provided an advantage (lower SRTs), compared with unilateral listening both in quiet and with the masker in front, but the variability within the group was substantial. In addition, children who use a CI and a hearing aid, on average, showed binaural “disruption” whereby the bilateral condition resulted in worse SRTs than the first-CI condition alone. This finding suggests lack of ability to integrate acoustic with electric hearing on this task. A disruptive effect of this nature has not been observed in all studies with bimodal children (e.g., Ching et al. 2005, 2006), suggesting that amplification approaches, residual hearing in the unimplanted ear, and possibly also the perceptual task may affect measures of bilateral benefits. In addition, it is important to recognize the fact that in many patients inputs at the two ears can be highly dissimilar, rendering comparison of

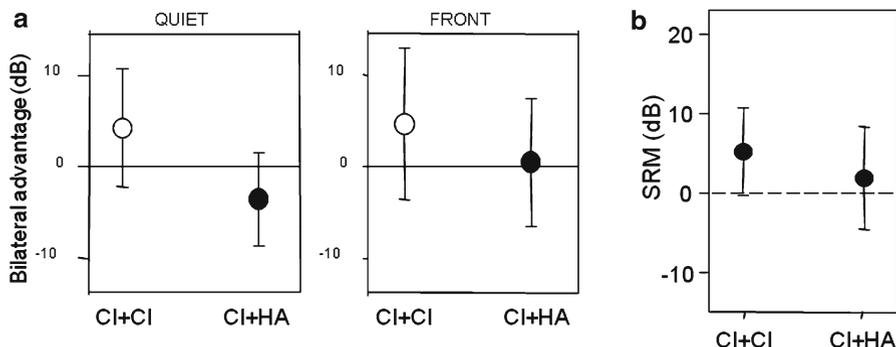


Fig. 6.9 Results from SRT measurements are compared for groups of children who were either fitted with bilateral CIs (CI+CI) or with a CI in one ear and a hearing aid (HA) in the nonimplanted ear (CI+HA). **(a)** Bilateral advantage is plotted for conditions in which target speech was presented from 0° front, and either no masker was present (*Quiet*) or the masker was at 0° front (*Front*). The Bilateral Advantage refers to the effect of using two devices versus using the CI in the first-implanted ear. Positive values indicate that adding the second CI or HA was beneficial relative to the first-CI alone. Negative values indicate that the CI+CI or CI+HA led to worse performance compared to the first-CI alone. **(b)** SRM (spatial release from masking) values are shown for the two groups of children. Positive values indicate that spatial separation of the target and masker was beneficial, whereas negative values indicate that spatial separation resulted in worse performance than in the collocated condition

time arrival or intensity difference at the two ears weak or absent. Therefore, improved speech intelligibility in noise may arise, not from binaural processing, but rather from monaural processing. That is, having two ears that each receive a signal provides listeners with advantages due to better signal/noise ratio in the ear that is farther from the noise.

Whereas the aforementioned studies focused on bilateral versus unilateral hearing and use of spatial cues, Grieco-Calub et al. (2009) were the first to examine the efficiency of language processing in toddlers who were profoundly deaf and had CIs compared with normal-hearing peers matched on chronological age. The approach was known as the looking-while-listening paradigm (e.g., Fernald et al. 1998). Children's eye movements (to a visual target matching a spoken word) were analyzed for reaction time, and accuracy. Children with CIs were less accurate and took longer to fixate target objects than age-matched normal-hearing peers. Both groups of children performed slower and less accurately in the presence of speech maskers that were collocated with the targets and presented at signal/noise ratio of +10 dB. An intriguing finding was that children with bilateral CIs did not perform any better than children with unilateral CIs, suggesting that the benefits of dual inputs may be restricted to conditions under which there is a poorer signal/noise ratio, or when spatial cues can be used to segregate the signal from maskers.

6 Conclusions

The binaural system of humans has a remarkable ability to process information in such a way that listeners become adept at localizing sounds, understanding speech in noisy situations, and suppressing echoes in reverberant spaces. In children and infants these abilities develop with time and experience. This chapter focused on research findings with young listeners who are normal hearing, and with their counterparts who are deaf and hard-of-hearing who use cochlear implants or hearing aids or both. In normal-hearing infants and children, rudimentary spatial hearing abilities appear during early infancy, and continue to develop and become more refined throughout childhood. In children with hearing impairment, a host of problems arise that can be attributed to factors such as the hardware and signal processing in the devices. In addition, it is important to recognize the fact that the auditory system of these young listeners is not completely healthy as it is likely to have undergone a period of deprivation before activation, and there exists loss of neural substrates that cannot be recovered. One of the threads that ties together the themes of this chapter is the importance of implementing tasks that are age appropriate that enable researchers to capture the perceptual effects of interest.

Acknowledgments The author is very grateful to students, postdoctoral fellows, and collaborators whose participation in the cited studies was important to the success of this chapter. They include Shelly Godar, Tina Grieco-Calub, Patti Johnstone, Soha Garadat, Smita Agrawal, Lynne Werner, and Jenny Saffran. The author received support for her work from the National Institutes of Health (R01 DC 003083 and 5R01 DC 008365).

References

- Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, 112(5 Pt 1), 2086–2098.
- Ashmead, D. H., Clifton, R. K., & Reese, E. P. (1986). Development of auditory localization in dogs: single source and precedence effect sounds. *Developmental Psychobiology*, 19(2), 91–103.
- Ashmead, D. H., Davis, D. L., Whalen, T., & Odom, R. D. (1991). Sound localization and sensitivity to interaural time differences in human infants. *Child Development*, 62(6), 1211–1226.
- Bernstein, L. R., & Trahiotis, C. (1992). Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies. *Journal of the Acoustical Society of America*, 91, 306–316.
- Bernstein, L. R., & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using “transposed stimuli.” *Journal of the Acoustical Society of America*, 112(3 Pt 1), 1026–1036.
- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*, revised edition. Cambridge, MA: MIT Press.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acustica*, 86, 117–128.
- Brungart, D. S., & Simpson, B. D. (2002). The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal. *Journal of the Acoustical Society of America*, 112(2), 664–676.

- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5), 975–979.
- Ching, T. Y., Hill, M., Brew, J., Incerti, P., Priolo, S., Rushbrook, E., & Forsythe L. (2005). The effect of auditory experience on speech perception, localization, and functional performance of children who use a cochlear implant and a hearing aid in opposite ears. *International Journal of Audiology*, 44(12), 677–690.
- Ching, T. Y., van Wanrooy, E., Hill, M., & Incerti, P. (2006). Performance in children with hearing aids or cochlear implants: Bilateral stimulation and binaural hearing. *International Journal of Audiology*, 45(Supplement 1), S108–112.
- Clifton, R. K., Morrongiello, B. A., Kulig, J. W., & Dowd, J. M. (1981). ‘Newborns’ orientation toward sound: Possible implications for cortical development. *Child Development*, 52, 833–838.
- Clifton, R. K., Perris, E., & Bullinger, A. (1991). Infants’ perception of auditory space. *Developmental Psychology*, 27, 187–197.
- Clifton, R., Freyman, R., Litovsky, R., & McCall, D. (1994). Listeners’ expectations about echoes can raise or lower echo threshold. *Journal of the Acoustical Society of America*, 95(5), 1525–1533.
- Cranford, J., Diamond, I. T., Ravizza, R., & Whitfield, L. C. (1971). Unilateral lesions of the auditory cortex and the ‘precedence effect.’ *Journal of Physiology*, 213(2), 24P–25P.
- Culling, J. F., & Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America*, 98, 785–797.
- Culling, J. F., Hawley, M. L., & Litovsky, R.Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *Journal of the Acoustical Society of America*, 116, 1057–1065.
- Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn Cunningham, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, 113, 2984–2987.
- Fernald, A., Pinto, J. P., Swingle, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9, 228–231.
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, 106, 3578–3588.
- Gabriel, K. J., & Colburn, H. S. (1981). Interaural correlation discrimination: I. Bandwidth and level dependence. *Journal of the Acoustical Society of America*, 69, 1394–1401.
- Garadat, S. N., & Litovsky, R.Y. (2007). Speech intelligibility in free field: Spatial unmasking in preschool children. *Journal of the Acoustical Society of America*, 121(2), 1047–1055.
- Godar, S. P., & Litovsky, R.Y. (2010). Experience with bilateral cochlear implants improves sound. *Otology Neurotology*, 31(8):1287–1292.
- Grantham, D. W., & Wightman, F. L. (1978). Detectability of varying interaural temporal differences. *Journal of the Acoustical Society of America*, 63(2), 511–523.
- Green, J. S., & Sanes, D. H. (2005). Early appearance of inhibitory input to the MNTB supports binaural processing during development. *Journal of Neurophysiology*, 94, 3826–3835.
- Grieco-Calub, T., & Litovsky, R. (2010). Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing. *Ear Hearing*, 31(5): 645–656.
- Grieco-Calub, T. M., Litovsky, R.Y., & Werner, L. A. (2008). Using the observer-based psychophysical procedure to assess localization acuity in toddlers who use bilateral cochlear implants. *Otology Neurotology*, 29(2), 235–239.
- Grieco-Calub, T., Saffran, J., & Litovsky, R.Y. (2009). Spoken word recognition in toddlers who use cochlear implants. *Journal of Speech Language Hearing Research*, 52(6), 1390–1400.
- Grose, J., Hall, J., & Dev, M. (1997). MLD in children: Effects of signal and masker bandwidths. *Journal of Speech Language Hearing Research*, 40, 955–959.
- Hall, J. W. 3rd, Buss, E., & Grose, J. H. (2007). The binaural temporal window in adults and children. *Journal of the Acoustical Society of America*, 121(1), 401–410.

- Hartmann, W. M. (1983). Localization of sound in rooms. *Journal of the Acoustical Society of America*, 74(5), 1380–1391.
- Hartmann, W. M., & Rakerd, B. (1989). On the minimum audible angle—a decision theory approach. *Journal of the Acoustical Society of America*, 85, 2031–2041.
- Hartung, K., & Trahiotis, C. (2001). Peripheral auditory processing and investigations of the “precedence effect” which utilize successive transient stimuli. *Journal of the Acoustical Society of America*, 110, 1505–1513.
- Hawley, M. L., Litovsky, R. Y., & Colburn, H. S. (1999). Speech intelligibility and localization in complex environments. *Journal of the Acoustical Society of America*, 105, 3436–3448.
- Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America*, 115, 833–843.
- Johnstone, P. M. (2006). Informational masking and spatial asymmetry in a “cocktail party” environment. University of Wisconsin-Madison, unpublished Ph.D. dissertation.
- Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type on speech intelligibility and spatial release from masking in children and adults. *Journal of the Acoustical Society of America*, 120(4), 2177–2189.
- Jones, G. L., & Litovsky, R. Y. (2008). Effects of uncertainty in a cocktail party environment in adults. *Journal of the Acoustical Society of America*, 124, 3818–3830.
- Joris, P., & Yin, T. C. (2007). A matter of time: Internal delays in binaural processing. *Trends in Neuroscience*, 30, 70–78.
- Kacelnik, O., Nodal, F. R., Parsons, C. H., & King, A. J. (2006). Training-induced plasticity of auditory localization in adult mammals. *PLoS Biology*, 4(4), e71.
- Kidd, G., Jr., Mason, C. R., Rohla, T. L., & Deliwala, P. S. (1998). Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*, 104(1), 422–431.
- Kollmeier, B., & Gilkey, R. (1990). Binaural forward and backward masking: Evidence for sluggishness in binaural detection. *Journal of the Acoustical Society of America*, 87, 1709–1719.
- Leibold, L. J., & Werner, L. A. (2006). Effect of masker-frequency variability on the detection performance of infants and adults. *Journal of the Acoustical Society of America*, 119(6), 3960–3970.
- Litovsky, R. (1997). Developmental changes in the precedence effect: Estimates of minimal audible angle. *Journal of the Acoustical Society of America*, 102, 1739–1745.
- Litovsky, R. Y. (1998). Physiological studies of the precedence effect in the inferior colliculus of the kitten. *Journal of the Acoustical Society of America*, 103, 3139–3152.
- Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*, 117(5), 3091–3099.
- Litovsky, R., & Ashmead, D. (1997). Developmental aspects of binaural and spatial hearing. In H. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing* (pp. 571–592). R Hillsdale, NJ: Lawrence Erlbaum.
- Litovsky, R. Y., & Clifton, R. K. (1992). Use of sound-pressure level in auditory distance discrimination by 6-month-old infants and adults. *Journal of the Acoustical Society of America*, 92(2), 794–802.
- Litovsky, R. Y., & Delgutte, B. (2002). Neural correlates of the precedence effect in the inferior colliculus: Effect of localization cues. *Journal of Neurophysiology*, 87, 976–994.
- Litovsky, R. Y., & Godar, S. P. (2010). Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. *Journal of the Acoustical Society of America*, 128(4), 1979–1991.
- Litovsky, R. Y. (2011). Spatial hearing abilities in 2-year old bilateral CI users. Invited presentation at the 13th International Symposium on Cochlear Implants in Children. Chicago, IL.
- Litovsky, R., & Macmillan, N. (1994). Minimum auditory angle for clicks with simulated echoes: Effects of azimuth and standard. *Journal of the Acoustical Society of America*, 96(2), 752–758.
- Litovsky, R. Y., & Madell, J. (2009). Bilateral cochlear implants in children. In L. Eisenberg (Ed.), *Clinical management of children with cochlear implants* (pp. 59–83). San Diego: Plural Publishing.

- Litovsky, R.Y., & Yin, T. C. (1998). Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. *Journal of Neurophysiology*, 80(3), 1285–1301.
- Litovsky, R.Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America*, 106, 1633–1654.
- Litovsky, R.Y., Johnstone, P. M., & Godar, S. P. (2006a). Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology*, 45(Supplement 1), S78–91.
- Litovsky, R.Y., Johnstone, P. M., Godar, S., Agrawal, S., Parkinson, A., Peters, R., & Lake, J. (2006b). Bilateral cochlear implants in children: Localization acuity measured with minimum audible angle. *Ear and Hearing*, 27, 43–59.
- Litovsky, R.Y., Parkinson, A., & Arcaroli, J. (2009). Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear and Hearing*, 30(4), 419–431.
- Litovsky, R.Y., Jones, G. L., Agrawal, S., & van Hoesel, R. (2010). Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *Journal of the Acoustical Society of America*, 127(1), 400–414.
- Litovsky, R.Y. (2011). Spatial hearing abilities in 2-year old bilateral CI users. Invited presentation at the 13th International Symposium on Cochlear Implants in Children. Chicago, IL.
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, 42, 135–159.
- Mills, A. (1958). On the minimum audible angle. *Journal of the Acoustical Society of America*, 30, 237–246.
- Misurelli, S. M., Zettler, C. M., Godar, S. P., & Litovsky, R.Y. (2010). Spatial release from masking in children: Symmetric and asymmetric masker distribution in the horizontal plane. Presented at the Meeting of the Association for Research in Otolaryngology, February 6–10, 2010, Anaheim, CA.
- Moore, J. M., Thompson, G., & Thompson, M. (1975). Auditory localization of infants as a function of reinforcement conditions. *Journal of Speech and Hearing Disorders*, 40, 29–34.
- Morrongiello, B. A. (1988). Infants' localization of sounds along the horizontal axis: Estimates of minimum audible angle. *Developmental Psychology*, 24, 8–13.
- Muir, D. W., Clifton, R. K., & Clarkson, M. G. (1989). The development of a human auditory localization response: A U-shaped function. *Canadian Journal of Psychology*, 43(2), 199–216.
- Nodal, F. R., Kacelnik, O., Bajo, V. M., Bizley, J. K., Moore, D. R., & King, A. J. (2010). Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets. *Journal of Neurophysiology*, 103(3), 1209–1225.
- Nozza, R. J., Wagner, E. F., & Crandell, M. A. (1988). Binaural release from masking for a speech sound in infants, preschoolers, and adults. *Journal of Speech and Hearing Research*, 31(2), 212–218.
- Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *Journal of the Acoustical Society of America*, 82(2), 454–464.
- Olsho, L.W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America*, 84(4), 1316–1324.
- Perris, E. E., & Clifton, R. K. (1988). Reaching in the dark toward sound as a measure of auditory localization in 7-month-old infants. *Infant Behavior and Development*, 11, 477–495.
- Peters, R., Litovsky, R. Y., Parkinson, A., & Lake, J. (2007). Importance of age and post-implantation experience on performance in children with sequential bilateral cochlear implants. *Otology & Neurotology*, 28(5), 649–657.
- Plomp, R., & Mimpfen, A. M. (1981). Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech reception threshold for sentences. *Acustica*, 48, 325–328.
- Pollack, I., & Pickett, J. M. (1958). Stereophonic listening and speech intelligibility against voice babble. *Journal of the Acoustical Society of America*, 30, 131–133.
- Populin, L. C. (2008). Human sound localization: Measurements in untrained, head-unrestrained subjects using gaze as a pointer. *Experimental Brain Research*, 190(1), 11–30.

- Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *Journal of the Acoustical Society of America*, 125(3), 1737–1743.
- Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, 61(3), 632–652.
- Svirskis, G., Kotak, V. C., Sanes, D. H., & Rinzel, J. (2002). Enhancement of signal-to-noise ratio and phase locking by a low threshold outward current in auditory neurons. *Journal of Neuroscience*, 22, 11019–11025.
- van de Par, S., & Kohlrausch, A. (1999). Dependence of binaural masking level difference on center frequency, masker bandwidth, and interaural parameters. *Journal of the Acoustical Society of America*, 106, 1940–1947.
- Van Deun, L., van Wieringen, A., Van den Bogaert, T., Scherf, F., Offeciers, F. E., Van de Heyning, P. H., Desloovere, C., Dhooge, I. J., Deggouj, N., De Raeve, L., & Wouters, J. (2009). Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. *Ear and Hearing*, 30, 178–190.
- van Hoesel, R. J. (2004). Exploring the benefits of bilateral cochlear implants. *Audiology & Neurotology*, 9, 234–246.
- van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of the Acoustical Society of America*, 113, 1617–1630.
- Werner, L. A. (1999). Forward masking among infant and adult listeners. *Journal of the Acoustical Society of America*, 105(4), 2445–2453.
- Whitfield, I. C., Cranford, J., Ravizza, R., & Diamond, I. T. (1972). Effects of unilateral ablation of auditory cortex in cat on complex sound localization. *Journal of Neurophysiology*, 35(5), 718–731.
- Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*, 118(5), 3164–3176.
- Wilmington, D., Gray, L., & Jahrsdoerfer, R. (1994). Binaural processing after corrected congenital unilateral conductive hearing loss. *Hearing Research*, 74(1–2), 99–114.
- Wright, B. A., & Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences of the USA*, 98(21), 12307–12312.
- Xia, J., Brughera, A., Colburn, H. S., & Shinn-Cunningham, B. (2010). Physiological and psychophysical modeling of the precedence effect. *Journal of the Association for Research in Otolaryngology*, 11(3), 495–513.
- Zurek, P. M., & Durlach, N. I. (1987). Masker-bandwidth dependence in homophasic and antiphase tone detection. *Journal of the Acoustical Society of America*, 81, 459–464.

Chapter 7

Development of Speech Perception

Robin Panneton and Rochelle Newman

1 Introduction

During infancy, interactions with caretakers provide young humans with opportunities to discover important aspects of their world, including ways in which they and others communicate. In the current view, “communication” involves dynamic, real-time processing of multiple sources of information arising both within and between participants, directing attention, enabling co-action, and enriching understanding of the surrounding world (Hollich et al. 2000; Locke 2001). Early human communication can take many forms, including the perception and production of music (e.g., maternal singing; see Trainor and Unrau, Chap. 8) and gesture, and clearly (for most children) involves their mastery of a native language system. The pathway to language in human infancy is not a quick one, but it is remarkably robust in that most children find their way to being fully communicative by the age of 3 years (Jusczyk 1997). That being said, there are important exceptions to this statement in that not all children reach the same level of language proficiency (Klee et al. 2000; see also Eisenberg et al., Chap. 9), with some showing marked deficiencies in communication by early toddlerhood, as in severe autism. In the end, developmental research must account for the full spectrum of functioning in this important domain. This is even more pertinent for practitioners who must bring resources to bear on children who are challenged in their abilities to communicate fully and effectively with others.

R. Panneton (✉)
Department of Psychology, Virginia Tech, Blacksburg, VA, USA
e-mail: panneton@vt.edu

R. Newman
Department of Hearing & Speech Sciences, University of Maryland, College Park, MD, USA
e-mail: newman1@umd.edu

This chapter provides an overview of the general time course for emerging skills related to the perception and production of speech across infancy and early childhood. The presentation begins with a brief description of methodologies used to assess the skills and deficits seen over the course of early language learning. Next, language learning is set within a motivational framework based on infants' early attentional focus; success in language learning critically depends on the ability of infants and young children to direct and maintain attention to sources providing information about their language systems. The ability to regulate infants' attention in language-related tasks is an essential aspect of research in this field. But beyond the laboratory, it is important to appreciate that language learning arises out of dynamic partnerships involving real-time adjustments in attention flow as relevant properties unfold and partners "communicate." Language learning occurs in such relationships. Subsequent sections summarize research that creates a portrait of the typical language-learning infant as one moving from a diffuse, superficial perceiver to a more highly focused and selective consumer of language. Last, the importance of attention regulation is emphasized in the context of challenges that infants face in natural language-learning conditions.

2 Measuring Speech Perception and Language Skills in Infancy

Infants are not always cooperative or accessible research participants. Nonetheless, great strides have been achieved in understanding early auditory psychophysics, perception, and language (Saffran et al. 2006), and in the ability of developmental scientists to probe infants for various abilities or limitations. Infants display an array of perceptual and cognitive skills that are themselves undergoing rapid development (e.g., statistical learning) and emerging in real time (i.e., participating in experimental protocols engenders learning in the moment). Developmental research on infants' perception of speech continues to advance in important ways as both methods and techniques evolve. Several excellent discussions are available to educate new workers in this field (McCardle et al. 2009; Johnson and Zamuner 2010), so the following is a selective summary of common practices and marked advances.

Research with young infants capitalizes on rudimentary motor responses (e.g., sucking on pacifiers, turning supported heads toward a sound) and responsive physiology (e.g., heart rate changes; cortical activation patterns). With age, infants become proficient at controlling their own body movements, allowing protocols to include volitional action (e.g., turning one's own head for specific consequences; Werker et al. 1997). Two common protocols for measuring speech perception involve either selective visual fixation of, or head-turn toward targets associated with speech (Johnson and Zamuner 2010). Visual fixation studies are designed to examine infants' speech preferences or speech discrimination. For example, some are designed such that fixation of a repeated visual target produces two kinds of speech streams, allowing investigators to gauge speech preferences by comparing

differences in looking times (Panneton-Cooper and Aslin 1990; Pegg et al. 1992). Other studies involve periods of familiarization (or habituation) during which fixation of a repeated visual target produces the same speech event, followed by discrete trials during which fixation produces either familiar or novel speech. Often, infants increase attention (i.e., look longer at the visual target) during novel trials, indicating speech discrimination (Fais et al. 2009; Sato et al. 2010).

Although the visual fixation technique works well across many ages, it is particularly well suited for younger infants given that it does not require head turns. Speech preference and discrimination studies with older infants are more likely to involve infants' head-turning to peripherally located lights, and involve some period of familiarization during which infants hear discrete ("cup") or fluent ("where is the cup?") speech. Next, infants hear both familiar and novel speech events and looking time is compared. This technique is typically referred to as the "head-turn preference procedure."

In spite of the utility in these behavioral techniques, three major methodological issues have surfaced (Aslin 2007). First, findings are typically group-level, and not always reflective of individual infants' performances. In the end, this compromises the predictive validity of research findings with respect to emerging language function, a problem for basic and clinical interests alike. Houston et al. (2007) recently addressed this concern by empirically validating a hybrid methodology for testing infants' speech discrimination by integrating robust elements from multiple protocols into one. After infants habituated to a nonsense word (e.g., "boodup"), they received a series of test trials including familiar ("boodup") and novel ("seepug") words, with two innovative design features. First, the ratio of novel to familiar trials was low (e.g., 3/14), increasing the saliency of novel presentations in the stream of test trials, as in classic "oddball" paradigms. Second, the two word types alternated during novel trials ("boodup/seepug/boodup/seepug...") decreasing the cognitive load for discrimination. Analyses of individual infants' data showed significantly higher rates of discrimination compared to other conditions including one feature but not the other. Moreover, these authors also found significant positive correlations between individual infants' performance across 2–3 days. Thus, the hybrid discrimination protocol promotes better internal as well as predictive validity.

A second methodological concern is the inconsistency of expected outcomes in terms of infants showing more attention to familiar versus novel information during test. Some infant speech perception studies find more attention to familiar than novel presentations, whereas other studies find more attention to novel than familiar presentations. Either outcome demonstrates discrimination, although whether attention is enhanced by the familiar or the novel affects conceptual interpretations (Houston-Price and Nakai 2004; Aslin and Fiser 2005). A more serious concern is when both patterns appear in a given sample (i.e., some infants prefer familiar speech whereas others prefer novel speech) leading to a null effect (and erroneous interpretations) at the group level, when discrimination was evident at the individual level (see Houston et al. 2007 for a discussion on this point). Multiple factors determine whether any given infant attends more to familiar or novel test events, such as stimulus complexity and individual processing strategies. Moreover, individual infants' attention to familiar versus novel information is related to concurrent and long-term measures of sustained attention and recognition memory (Colombo 2001).

Third, most protocols used to investigate infants' speech perception do not provide graded responses. For example, in a segmentation task, the primary measure of interest is whether infants attend longer on familiar or novel trials (an either/or measure). Typically, there is no quantification of response strength; even if all infants look more at the familiar event, do some infants look longer (or faster) than others, or show a stronger effect in one task than another? Estimating strength of preference or discrimination could clarify whether some cues to segmentation make the task easier or harder than others. Improved measurement may emerge from advances in using eye-tracking systems to assess the speed, direction, and duration of infants' fixations on visual targets that have been associated with verbal labels (e.g., McMurray and Aslin 2004; Fernald et al. 2008).

A complementary approach to the use of behavioral protocols for understanding speech perception in infants is provided by physiological techniques (e.g., autonomic responses such as heart rate or central responses such as brain activity), some of which can stand alone as measures of processing or work in conjunction with behavioral tasks. Researchers also use different scalp-level recording methods to specify cortical and subcortical involvement in early language processing (Friederici and Oberecker 2008). Brain-relevant recording procedures continue to be refined for infants and young children, such as scalp electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG). One emerging technology involves the use of near-infrared spectroscopy (NIRS), a technique measuring changes in hemodynamic flow in cortex as a function of event exposure (Mehler et al. 2008; Lloyd-Fox et al. 2010). The appeal of NIRS stems in part from its greater cortical source precision compared to other scalp recording systems, and from the fact that it can be used in awake, alert infants (Aslin and Mehler 2005). Recent studies employing fMRI or NIRS with infants across a wide age span have shown increased responding in temporal cortex to voices (e.g., superior temporal sulcus), either presented alone or accompanied by faces, along with patterns of lateralization that reveal the early specialization of information processing in development (Dehaene-Lambertz et al. 2010; Grossmann et al. 2010). Perhaps the need for the aforementioned graded measurements will be achieved by coupling the use of behavioral tasks with psychophysiology (e.g., heart-rate change, amplitudes of deflections in ERP waveforms; see Junge et al. 2010).

In the meantime, developmental studies on speech perception depend importantly on cross-laboratory confirmation of both positive and negative findings related to emerging skills and how they correspond to later levels of language functioning. Although this kind of collaborative validity is ideal, some important questions remain unverified when participant groups are harder to recruit or techniques are more difficult. A key element in the success of developmental studies with infants is eliciting and maintaining task-related attention. Laboratory studies (c.f. Kuhl et al. 2003; Goldstein and Schwade 2008) demonstrate that infant learning is dependent on gaining infants' attention. Even more so, language learning outside of the laboratory demands recruiting and sustaining attention in situations that involve multiple sources of information. For that reason, the next section of the chapter offers a discussion of the context within which language learning occurs during infancy, both in and outside the laboratory.

3 Attention to Language Function in Infancy

One robust finding in the early language-learning literature is that caretakers around the world modify their communicative style when addressing infants. These modifications are dynamic, in that they are tuned both to the sociolinguistic development of the child and to the context in which the interaction occurs (Panneton-Cooper 1993; Kitamura and Burnham 2003). Importantly, some aspects of these changes recruit and maintain attention to language, occurring not only in caretakers' vocal acoustics, but in their facial expressions and gestures as well.

Infant-directed speech (IDS) is distinguished acoustically by higher pitch (fundamental frequency or F0), more exaggerated pitch contours, larger pitch range (the difference between F0 maximum and minimum), slower tempo, longer pauses, and higher rhythmicity than adult-directed speech (ADS; Fernald and Mazzie 1991; Katz et al. 1996). Such prosodic exaggeration most likely scaffolds acquisition of linguistic structure in infants as it appears to do in adults (Golinkoff and Alioto 1995). Developmentally, even newborn infants prefer IDS when the alternative is ADS (Panneton-Cooper and Aslin 1990; Pegg et al. 1992). In addition, newborns prefer recordings of their mothers' voices compared to those of unfamiliar females (DeCasper and Fifer 1980), with evidence that such learning is influenced by prenatal experience (DeCasper and Spence 1986; Mehler et al. 1988). Integrating across these early biases leads to the prediction that newborns' attention would be maximally heightened by recordings of maternal IDS. However, young infants (1-month-olds) do not prefer IDS to ADS when both are spoken by their own mothers, although this preference is evident if *nonmaternal* recordings are used (Panneton-Cooper et al. 1997). This finding supports the view that the recruitment of infants' attention to language is heightened by whatever context is most familiar and meaningful to infants at that developmental time. For newborns, the maternal voice is a primary attractor; with more experience, preference for the maternal voice becomes refined to include the kinds of acoustic exaggerations that are shown in caretaking episodes around the world (Fernald et al. 1989). In fact, 4-month-olds do prefer maternal IDS over maternal ADS (Panneton-Cooper et al. 1997) as well as IDS over ADS when spoken by unfamiliar females (Fernald 1985), suggesting that early in the first year, infant–mother exchanges promote the extension of infants' selective attention to speakers and speaking style.

Differential attention to IDS persists at least until the last quarter of the first postnatal year with the acoustic and lexical characteristics of IDS shifting as infants' abilities improve and parental communicative intentions become more complex (Kitamura and Burnham 2003). However, the ability of vocal prosody alone to promote infants' attention appears to diminish with age (Hayashi et al. 2001; Newman and Hussain 2006), suggesting that the contribution of IDS to language learning is more evident in early rather than later infancy.

Early on, IDS most likely garners its perceptual pull in moderating infants' attention from its association with vocal emotion (Kitamura and Burnham 1998; Singh et al. 2002) rather than from the fact that it is speech to infants per se. Several researchers (e.g., Trainor et al. 2000; Santesso et al. 2007) argue that infants are drawn toward

IDS primarily because of its exaggerated emotional tone, which then supports language acquisition. As such, research on IDS has shifted from an analysis of acoustic properties to the affective components that underlie infant preferences. Exaggerated IDS prosody typically reflects vocal expression of emotion, such that the ability of IDS to regulate infant attention may stem from emotional engagement (Kitamura and Burnham 1998). Infants prefer positive over negative vocal affect at least through the first half of the first year (Panneton et al. 2006). Infants also discriminate specific categories of IDS at different points during development, with those signaling emotional intent more preferred by younger, but not by older infants (Moore et al. 1997; Kitamura and Lam 2009). In the adult literature, emotional tone of voice enhances processing of lexically ambiguous words (Nygaard and Lunders 2002) and improves word recognition memory (Dietrich et al. 2000), even though the necessary and sufficient acoustic correlates of vocal emotion that aid perception remain unclear (Scherer 1995).

Importantly, the ability of positive-emotion speech to increase infants' attention is contingent on the nature and quality of exchanges between infant and caretaker. As is clear from studies looking at infant–caretaker dynamics (Stern 1985), early communication is bidirectional in nature, and the ability to increase and maintain infant attention is compromised when contingency is diminished, even in the face of positive emotion in voice and other gestures. An interesting study in this regard involved recording mothers' speech to their infants while they both interacted over a monitor (i.e., the mother and infant were in separate rooms). Mothers were asked to use their voices to increase positive emotion in their infants, only unbeknownst to the mothers the infants could not hear them. A confederate female either engaged the infants positively or negatively, whenever their mothers spoke. Acoustic analyses of the mothers' voices indicated significant elevation of pitch and pitch variance in the group whose infants received surreptitious positive regard, supporting the notion that vocal adjustments in caretakers' speech are highly influenced by contingent adjustments in infant behaviors during interaction (Smith and Trainor 2008). Thus, infants' primary motivation for attending to information in the speech stream arises out of the dynamic exchange known as the infant-directed context. So what is learned in these ongoing interactions?

4 Attention to Language Structure in Infancy and Early Childhood

Within the early context of heightened attention to IDS, perceptual shaping of attention to language takes root. One clear benefit of infants' heightened attention to IDS is access to important aspects of native language structure (Fisher and Tokura 1996; Christophe et al. 2003). Considerable research effort has focused on how IDS bootstraps infants' early lexical and syntactic awareness. Prosodic information plays a vital role even in adult speech processing, especially in English where intonational emphasis is used to accent novel information (Gerken 1996). Boundaries of both

sentences and restrictive relative clauses in English are marked by falling pitch contours and final vowel lengthening, and words (even syllables) conveying new information are higher pitched and longer. Infants take advantage of prosody in learning how to segment and categorize speech input (Gerken et al. 2005; Thiessen et al. 2005), and in syntactic acquisition (Soderstrom et al. 2003).

4.1 In the Beginning: Perceptual Shaping of Attention to Speech in the First Six Months

In general, good empirical support exists for the ability of young infants to attend to and recognize various aspects of human speech (e.g., preferences for the native language in newborns, Moon et al. 1993), prompting focus on aspects of languages that lend to their uniqueness. Languages differ from one another in a number of ways, but many of those differences are in suprasegmental properties, particularly in terms of prosody and rhythm, and these seem particularly critical for infant language preferences. Languages can be classified into three main categories on the basis of their general rhythmic structure (Pike 1945; Abercrombie 1967), providing one potential cue to infants as to native versus nonnative designation. Languages are viewed as “syllable-timed” in which each syllable has an equivalent duration in production, “stress-timed” in which time between stressed syllables is more-or-less constant and unstressed syllables are shortened to fit between the stressed-syllable beats, or “morae-timed” in which syllables consist of either one or two subsyllabic durational units, called morae. Classifying languages into rhythmic groups is now viewed as overly simplistic given that languages exist along a continuum of timing patterns, and languages within a “category” may still differ from one another rhythmically. Nonetheless, this basic rhythmic distinction is readily perceived by adults, and even by nonhuman primates when most of the segmental information has been removed (Ramus et al. 2000). Infants’ familiarity with a particular language rhythm may enhance discrimination and provide an early basis for native language preferences.

As mentioned previously, newborn humans prefer their native language by as young as 2 days of age (Mehler et al. 1988; Moon et al. 1993), demonstrating the ability to distinguish between languages, and appear particularly sensitive to language rhythm. Both 5-month-old infants and newborns discriminate two unknown languages when they fall into different rhythmic classes (Nazzi et al. 1998), but fail to do so when the languages come from the same rhythmic class (e.g., English vs. Dutch), unless one is native (Nazzi et al. 2000). However, Christophe and Morton (1998) found that 2-month-olds did not discriminate languages from different rhythm classes, either indicating a U-shaped developmental trend, or resulting from methodological differences across studies.

Thus, discriminating languages within a rhythm class depends on whether infants are highly familiar with one. Similarly, infants discriminate two dialectal versions of their native language, but only if one of the dialects is familiar (Nazzi et al. 2000; Butler et al. 2010). It is not clear how this process works. One possibility is that

infants gradually refine discrimination within a rhythmic class, moving from reliance on gross prosodic differences to more fine-grained differences, based on experience with their native language (Nazzi et al. 2000). However, interpreting these patterns of results is complicated by the methodological differences across studies and the lack of longitudinal approaches in this area of work. Another way to address this issue developmentally is to examine bilingual infants, given the demand on these infants to discriminate two native systems. Although scarce, available results suggest that bilingual infants discriminate their two languages by 5 months, even when the language systems fall into the same rhythmic class (Bosch and Sebastián-Gallés 2001). How consistent this discrimination ability is across infants or across languages has not been studied, despite frequent concerns from parents regarding the benefits versus risks of raising children bilingually. Interestingly, one recent study found differences between mono- and bilingual infants' discrimination of native from nonnative *visual* speech, an area of interest given that prosody and rhythm in language can not only be heard, but can also be seen (Munhall et al. 2004). As speakers engage in conversation, they move their facial muscles, heads, and bodies in ways that spatiotemporally correlate with their speech. Weikum et al. (2007) found that 4- and 6-month-old English-learning infants discriminated silent videotapes of women speaking English versus French, suggesting that infants attune to how speech is *visually* represented in human facial movement. In contrast, 8-month-olds did not make this discrimination across visual languages unless they were English–French bilinguals.

Although newborns' preference for natural recordings of their native language may be primarily based on rhythm (Moon et al. 1993), newborns also show preferences for canonical consonant-vowel compounds (compared to synthetic analogs; Vouloumanos and Werker 2007). Extending to slightly older infants, Panneton-Cooper and Aslin (1994) also found preferences for normal IDS speech over sine-wave analogs of IDS in 4-month-olds. These preferences may reflect early bias toward natural speech, and are supported by neurophysiological studies on early sensitivity to language-specific information (Peña et al. 2003). This early perceptual shaping of preferences for suprasegmental aspects of native languages is complemented by attention to more fine-grained segments, such as consonants and vowels. Perception of microstructure in speech led to early studies examining infants' categorical perception of phonemes. Categorical perception is seen when listeners are poor at distinguishing sounds from the same phonetic category (e.g., different tokens of the sound “b”) but successful at distinguishing sounds from different phonetic categories (e.g., “b” from “d”). Categorical perception poses a puzzle for development because phonetic categories are language specific; the same acoustic difference could occur within a category in one language (and be ignored by native adult listeners) but could signal an important phonetic distinction in another language. Thus, categorical perception cannot be fully “set” from birth, implying that during development infant listeners must either learn to distinguish phonetic categories or learn to ignore the differences within categories.

Remarkably, categorical perception of phonemes appears early in infancy. Eimas et al. (1971) presented 1- and 4-month-old infants with a synthetic CV syllable,

either /ba/ or /pa/. After the infants' attention habituated to the item, infants either heard further examples of the same item (control group), a switch to a new token within the same category, or a switch to a token from the opposite category. Critically, the two "switch" tokens were equivalently dissimilar acoustically from the original item. Infants' attention showed significantly greater recovery for the item in the opposite category, indicating greater discrimination of items across a category boundary than within a category. Follow-up studies reported similar results for a range of phoneme distinctions, in a range of syllabic contexts (e.g., Eimas and Miller 1980; Cohen et al. 1992). Importantly, these studies also showed that infants discriminated contrasts that did not occur in their native language, contrasts adult speakers of the same language failed to discriminate (e.g., Trehub 1976). Infants also failed to distinguish some contrasts that were within their native language (Lasky et al. 1975). Thus infants' phoneme discrimination appears to be language-universal at first, becoming increasingly attuned to the native system as infants gain experience with hearing speech.

4.2 Second Half of the First Year: Perceptual Attunement and Attentional Pruning

In the first 6 months of life, infants focus their attention on rhythmic and prosodic aspects of language. In contrast, their initial perception of segmental properties appears to be less colored by their native language. Comparatively, attention pruning dominates speech perception during the next 6 months, with infants attending more selectively to finer-grained aspects of their native language. Progressive attunement to experientially dependent aspects of information across infancy may extend to other domains as well (e.g., face processing; Scott et al. 2007).

Werker and Tees (1984) explored the time course of increasing attunement to native phonemes across infancy. They tested English-learning infants ages 6–8 months, 8–10 months, and 10–12 months on two nonnative contrasts, from Hindi and Thompson/Nlaka'pamux, that adult English-speakers failed to distinguish. The youngest infants discriminated both contrasts, but older infants failed to do so, suggesting that there was a shift in sensitivity to nonnative phonemes between 8 and 10 months of age. However, not all nonnative contrasts present this difficulty for older infants. Best et al. (1988) found no attenuation of discrimination in English infants for Zulu clicks (see also McMurray and Aslin 2005). Nonetheless, changes in discrimination have been found at roughly the same age for a variety of segmental distinctions (Polka and Werker 1994; Tsao et al. 2006), including both declines in the discrimination of nonnative contrasts, and improvements in, or in some cases development of, discrimination of native contrasts (Hoonhorst et al. 2009; Sato et al. 2010).

Languages differ not only in their phonemic patterns, but also in their use of vocal tone. According to some estimates, half the people in the world (and perhaps as many as 70%) speak a tone language (Yip 2002), in which differences in pitch shape (i.e., fundamental frequency) serve to distinguish meanings, much the same

way phonetic differences do. Currently, little is known about how tonal languages are perceived by infants (Yip 2002). Studies of infants learning both Chinese (Mattock and Burnham 2006) and Yorùbá, a tonal language of west Africa (Harrison 2000) found that infants respond to changes in tone across syllables that are phonemic in function, and appear to do so in a roughly categorical manner. In contrast, infants learning English show early discrimination of tonemes, but reduced discrimination by 9 months of age (Mattock et al. 2008), corroborating similar findings that infants' attention to nonnative segmental distinctions diminishes as experience with the native language increases.

Thus, there is strong evidence for some form of attunement to one's native language during the second half of the first postnatal year (often referred to as "perceptual reorganization"). As infants gain more experience with their native language, their perception (and production; Boysson-Bardies et al. 1984) both change, molding themselves to the typical input. This learning takes place within the attention framework of IDS, in that IDS often consists of elongated and more clearly enunciated vowels (e.g., Kuhl et al. 1997), and the degree of hyperarticulation (i.e., vocal clarity) present in mothers' speech positively correlates with infants' performance on a phoneme discrimination task as well as their vocabulary growth (Liu et al. 2003). Exaggerated pitch contours alone appear to facilitate vowel discrimination in 6–7-month-olds (Trainor and Desjardins 2002). Thus infants' inclination to attend selectively to IDS may enhance their learning of the segmental properties of fluent speech, which then leads to better attention to the input, and better subsequent learning of other important components of their native language.

What of phonemic attunement in infants raised in multilingual environments? Bosch and Sebastián-Gallés (2003) compared monolingual Spanish, monolingual Catalan, and bilingual Spanish/Catalan infants on a phonetic contrast (/e-eh/) that occurs in Catalan, but not Spanish. Four-month-olds discriminated the contrast regardless of which language(s) they learned in the home, but by 8 months only monolingual Catalan infants made the distinction. Diminished discrimination was expected in the monolingual Spanish infants, but surprising in the bilingual infants, for whom the contrast was still relevant. In a follow-up study, 12-month-old bilingual infants regained discrimination of the contrast, suggesting that the time course for attunement might be different in infants learning more than one language (see also Sebastián-Gallés and Bosch 2009). In contrast, other studies with bilingual infants show a time frame for language-specific phonetic reorganization similar to that for monolinguals (Burns et al. 2007; Sundara et al. 2008). It is unclear whether these inconsistencies are the result of the methods used, the specific languages or phonetic contrasts tested, or some other difference. Sundara and Scutellaro (2010) suggest that infants learning two rhythmically different languages may perform more similarly to monolinguals than do infants learning two rhythmically similar languages, perhaps because infants can use rhythm as a means of sorting the input, thus avoiding confusion. However, additional work comparing monolingual and bilingual populations is needed, particularly given the high rate of bilingualism across the world.

The mechanism for phonemic attunement derives from infants' perception of whether a particular sound distinction makes a meaningful difference in the

language, rather than being based on whether the sound distinction is heard in the ambient surroundings. For example, many English speakers produce prevoicing in their stop consonants, but this does not affect the interpretation of the words, and thus prevoiced stop consonants are not treated as a separate sound category. So how do infants learn which sounds are linguistically important? This is perplexing given that native attunement occurs when infants do not yet know many words, and thus could not know which phonetic distinctions make lexically-important distinctions in their language. Maye et al. (2002) suggest that the statistical distribution of sounds provides information as to their importance. More specifically, if variation along some acoustic measure was important in the language, the distribution of sounds would form a bimodal pattern: speakers would avoid producing tokens that fell at the category boundary, and productions would instead cluster around two (or more) distinct category centers or prototypes. In contrast, if variation along a given acoustic measure was irrelevant, speakers' productions would form a unimodal distribution. Maye et al. (2002) first familiarized two separate groups of infants with speech sounds from either a unimodal or bimodal distribution, then tested infants' discrimination of tokens along the entire continuum. Only the infants who had been familiarized with a bimodal distribution of sounds discriminated changes in speech tokens, suggesting that infants track the distribution of sounds they hear. This kind of sensitivity to distributional properties of speech provides a potential mechanism whereby attunement could occur.

Interestingly, distributional learning of phoneme categories in infants appears to also be influenced by visual speech information (Teinonen et al. 2008). Given that distributional cues for phonetic categories occur in the input (Werker et al. 2007) and that these are learnable by computer models (Vallabha et al. 2007), they could presumably be learned by infants as well. That said, it remains unclear whether speech input to infants is consistent in this respect for all phoneme categories and for suprasegmental distinctions, such as tone categories. Determining the number of tone categories in a given language is a critical prerequisite for learning the language, and determining whether the distributional properties would signal this to infants is of particular importance. It is also unclear how well infants' ability to track such distributions in a laboratory setting will "scale up" to more real-world settings, or whether there might be individual differences in this ability.

In addition to learning the phonetic structure of their native language, infants also learn *patterns* of phoneme distributions, referred to as phonotactics, toward the end of their first postnatal year. Jusczyk et al. (1994) showed that 9-month-olds, but not 6-month-olds, listened longer to lists of items that had more common phonetic patterns than to lists with less common patterns. However, the high-probability sequences contained both high positional phoneme frequency (that is, the segments were common in that word position), and contained more common phoneme combinations (or biphones) and higher-probability phonemes. Thus, there were multiple forms of phonotactic information that infants could have perceived from the input, and it is not clear which drove infants' preferences, or whether infants are equally sensitive to them at this age. In addition, the items in this study were entirely CVC tokens; thus aspects of phonotactics having to do with consonant clusters were not

investigated. Languages differ substantially in the number and types of consonant clusters allowed, from languages that forbid all consonant clusters within a syllable (such as Japanese) to languages that allow large strings (such as Slovak in which there is an entire tongue twister without vowels, *strc prst skrz krk*; Hanulíková et al. 2010), and infants might be expected to pick up on this aspect of phonotactics quite early. Although infants become aware of phonotactic probabilities during their first year of life, the specifics of how this process occurs, or how gradual this acquisition might be is unknown; it seems likely that the acquisition of phonotactic biases takes place along a continuum, with some aspects being recognized earlier than others, and there may be limits on which statistics infants will track. Little research has attempted to compare different types of statistical patterns, or to examine how infants determine which statistical computations might be relevant (Soderstrom et al. 2009).

Finally, it is not clear *how* infants track such patterns. One possibility is that infants store patterns of input, comparing new input to the combined set of prior exemplars. Preferences for more common statistical patterns would arise out of the process of recognition, with infants storing large amounts of relatively unprocessed data. An alternative approach is that infants track patterns during their original perception, and store these outcomes rather than the raw data. Certain statistical patterns would be more likely to be observed than others, since untracked properties could not emerge at a later date. This distinction is akin to that of prototype versus exemplars in categorization, but has received less attention in the infant language literature (c.f. Polka and Bohn 2011 for an excellent conceptualization of this prototype model for speech perception).

Although segmental changes may be evidence of a focusing of infant attention, the underlying goal of communication is not to discriminate phonemes but instead to understand the speaker's intention, which requires recognizing meaningful units (words). Most of the speech that infants hear consists of multiword utterances without obvious pauses or breaks demarcating boundaries; learning from this requires that infants first separate the fluent speech into individual words, a task referred to as word segmentation. Jusczyk and Aslin (1995) developed the first paradigm for evaluating the development of this segmentation ability. Using the head-turn preference procedure discussed earlier, these authors familiarized infants with two target words (either *cup* and *dog*, or *bike* and *feet*), spoken in isolation. Next, infants were presented with four fluent speech passages, each of which contained one of the potential target words. Infants age 7.5 months, but not those age 6 months, listened longer to the passages containing familiarized words, showing evidence of segmentation.

More recent work has shown that the precise age at which segmentation first occurs depends on both the method of testing (e.g., ERP results demonstrate slightly earlier segmentation ability than that shown behaviorally; Kooijman et al. 2005), and the type of lexical unit (words beginning with vowels, words with atypical stress patterns, and verbs show slightly later segmentation abilities; Mattys and Jusczyk 2001; Nazzi et al. 2005). However, across these studies, the ability to segment is consistently shown to develop as infancy advances. Thus, although the exact age varies, the general pattern is consistent—segmentation precedes most types of

word learning. Moreover, infants who show stronger segmentation skills also show enhanced language-learning skills at later ages (Newman et al. 2006; Junge et al. 2010), supporting the notion that this ability to segment may be a critical skill underpinning language acquisition.

As older infants' skills at segmentation increase, they begin to rely on a greater variety of cues to segment successfully. One knowledge-based cue promoting segmentation is word familiarity (sometimes referred to as *segmentation by lexical subtraction*, as in White et al. 2010). Words that are highly familiar to infants promote segmentation of adjacent words, even if those are less familiar. For example, infants can segment words adjacent to the word "mommy" (*Mommy's feet* vs. *Lola's dog* produces better segmentation of *feet*; Bortfeld et al. 2005). Another knowledge-based cue to segmentation is infants' ability to use statistical patterns in the spoken input. Certain phonemes are more likely to occur in particular word positions (e.g., "ng" in English only occurs syllable-finally), and certain pairs of syllables are more likely to occur together (as part of the same word) than others (e.g., /nt/ is more likely to occur subsequent to /In/ than as the start of a new word). Computer models tracking such probabilistic patterns consistently identify word boundaries (Brent and Cartwright 1996; Christiansen et al. 1998). Importantly, Saffran et al. (1996) demonstrated that infants use statistical cues in segmentation; after hearing a 2-min stream of continuous speech in which some syllables were adjacent to one another more regularly than were others, infants listened longer to atypical syllable combinations (those that had occurred adjacently less frequently) than to more common syllable combinations. This suggests infants treat high-probability sequences as potential new words. Interestingly, infants are more successful at linking these high-probability sequences with objects than they are at linking low-probability sequences with objects (Graf Estes et al. 2007).

Infants also use a wide variety of acoustic cues in segmentation, such as prosodic or stress cues, coarticulatory cues, phonotactic cues, allophonic cues, and cues to phonological phrase boundaries (see Saffran et al. 2006, for a review), although none of these cues provides definitive information in all settings. Another aid to segmentation is word position within an utterance, as 8-month-old infants are better able to segment utterance-initial or utterance-final words than utterance-medial words (Seidl and Johnson 2006). Overall, infants perform better at segmentation tasks when words have been "partially segmented" for them, and the process of segmentation appears to be one of integrating a wide array of probabilistic sources of information.

Given that segmentation cues appear to become available to infants at different stages in development, infants' weighting of potential cues also changes as they gain more experience with their native language (e.g., Johnson and Jusczyk 2001; Thiessen and Saffran 2003). In general, infants appear to move from depending on syllabic properties, such as statistical regularities of syllables and lexical stress, to being able to use more detailed segmental information, in the form of phonotactic and allophonic cues. Mattys et al. (2005) proposed one detailed cue hierarchy, but arguments over the relative weightings of different cues remain and these could vary across languages. As infants become more efficient in their processing, they are able

to integrate more types of information simultaneously (Morgan and Saffran 1995). As a result of these changes, the ability to segment the speech stream has a drawn-out developmental time course.

There are several important issues about the development of segmentation skills that remain virtually unaddressed. First, the vast majority of research on segmentation has focused on either English or other Germanic languages. Cross-linguistic differences in syllabic structure or in the use of affixes is likely to influence segmentation strategies, but only a few studies have explored segmentation in other European languages, let alone languages that may be more dissimilar either acoustically or structurally. Second, the literature on early segmentation has arisen from the study of infants raised in a monolingual environment. Many of the cues for segmentation, such as stress patterns, are language specific, and segmentation processes may be quite different for infants learning multiple languages simultaneously. Third, most of the research on segmentation has focused on finding the youngest age at which infants could reliably succeed on a task. Presumably, the ability to segment speech may not be a skill that infants either have or do not have, but may instead be a skill that they continue to develop over an extended time frame, and far less work has explored this developmental progression. Finally, there are a few studies that have suggested that infants who demonstrate more advanced segmentation skills are likely to show enhanced language-learning skills at later ages, perhaps because their early segmentation ability has provided them with more opportunities to learn words and morphemes. Although segmentation has been shown to be delayed in children with Williams syndrome (Nazzi et al. 2003), most research has focused on variation within a typically developing cohort; testing segmentation abilities in at-risk populations will be particularly fruitful for the early identification of language delay or disorder.

4.3 Is Speech Perception in Infancy Related to Emerging Language Proficiency after the First Year?

As is evident in Sects. 4.1 and 4.2, a great deal of research has focused on infants' perception of various cues to language structure between 6 and 12 months of age (i.e., parsing/segmenting words sensitivity to conditional probability of adjacent units). Clearly, infants are progressively attuning perception to speech information such that they are primed for processing within native-language contexts (Kuhl 2007). As a result, one would predict older infants and toddlers would show impressive competencies in their initial grammatical constructions (e.g., rich and flexible surface structure) as they begin to communicate. Although pattern recognition is a first step in language learning, toddlers who attempt to communicate with others need to invoke such patterns in context, and construct utterances with them in ways that accurately reflect intention and meaning. In the process of communicating, one can afford to minimize and even omit a fair amount of structure as long as meaning and intention are preserved. The important question at stake for those interested in

language trajectories, then, is whether there is any demonstrable link between the perceptual acumen of infants and their subsequent communicative skills.

One interesting paradox between the literatures on speech perception in infancy and lexical/grammatical development in toddlerhood is that toddlers show lower levels of grammatical/syntactic sophistication than expected. That is, the productive strategies that characterize toddlers' first attempts at communicating do not seem to include many of the lexical units or relations between units that they perceived during infancy. Children's vocabulary certainly makes rapid advances after the first postnatal year, but it is not clear how infants' perception of form and function translate into their own productive strategies later on. At least one study has found promising predictive validity with the head-turn preference procedure. Newman et al. (2006) analyzed the relationship between infants' segmentation of words from fluent speech and subsequent measures of language proficiency (e.g., vocabulary size). Aggregate data from several studies showed that after familiarizing 7.5- to 8-month-olds to single words, infants (as a group) attended more to sentences that contained those same words, than to those composed of novel words. Importantly, a longitudinal follow-up of individual infants from these studies revealed that infants who performed *poorly* in this word segmentation task also showed significantly lower vocabulary sizes at 24 months.

Although a complete treatment of this paradox is beyond the scope of the current chapter, readers are directed to a thoughtful discussion by Naigles (2002), in which she argues that speech perception in infancy has much less to do with *meaning* than does the active utterance construction by young children. That is, infants operate on continuous speech from the perspective of naïve pattern recognition (e.g., adjacent phonemes with higher conditional probabilities are recognized, processed, and retrieved as clusters), rather than from the perspective of what a given cluster actually denotes about an object or object relations. An excellent example of this comes from studies using the "switch paradigm" to study object–word association learning. In one case, infants at age 14 months readily learn the associations between ObjectA-LabelA and ObjectB-LabelB when the objects and labels are novel, and when the labels are maximally contrastive (e.g., "neem" vs. "lif"; Werker et al. 1998). However, same-age infants do not show evidence of object–word association learning when the labels involve minimal-pair distinctions (i.e., "bin" vs. "din"; Stager and Werker 1997; c.f. Fennell and Werker 2003), even though younger infants have no difficulty discriminating "bin" from "din." Thus, although infants at this age can distinguish minimal pair contrasts, the increased cognitive demand of object–word pairing makes the minimal pair contrast harder to process. In this example, knowing something about younger infants' perceptual acuity does not lead to accurate predictions about their later ability to apply this skill in a learning context.

As speech perception and discrimination abilities continue to evolve, toddlers make significant gains in information processing during their 2nd/3rd years. Bernhardt et al. (2007) found significant, positive correlations between 14-month-olds' word–object association learning and measures of their expressive and productive language skills up to 2 years later. Two-year-olds, compared to 15-month-olds, more accurately and rapidly locate named objects without even hearing entire verbal

labels, or when the labels themselves are acoustically degraded (Fernald et al. 2001; Zangl et al. 2005). Such processing efficiency correlates positively with better lexical and grammatical skill over age, not only in absolute performance, but in growth trajectories as well (Fernald et al. 2006). Moreover, speed of spoken word recognition and vocabulary size during toddlerhood correlate positively with multiple indices of linguistic skill (e.g., expressive vocabulary, formulating sentences, word structure) and working memory performance at 8 years of age (Marchman and Fernald 2008). Thus, the speed with which older infants and toddlers process continuous speech is one factor that plays an important role in emerging linguistic competency across childhood.

Nonetheless, an important tension exists between being fast and being accurate in the continuous perception and production of speech. Other studies have concentrated on identifying mediating factors for toddlers' comprehension of spoken words. For example, 11-month-olds recognize familiar words, but only if their onset and offset phonemes remain intact [e.g., infants recognize the familiar word "dirty" but do not recognize the similar nonwords "nirty" (onset violation) or "dirny" (offset violation); Vihman et al. 2004; Swingley 2005]. Such sensitivity to perceptually salient portions of words indicates that even as toddlers increase their speed of processing, violations of expected phonotactic sequences can attenuate comprehension.

Graf Estes et al. (2007) found that 17-month-olds were more likely to learn object-label associations from words that they had previously segmented from running speech compared to nonwords. Similarly, 2-year-olds use a combination of familiarity and prosody (e.g., word stress) to disambiguate function in sentences, such that their ability to accurately locate a picture of a familiar noun ("doggy") was unaffected by preceding *unstressed* adjectives, whether familiar ("good doggy") or unfamiliar ("glib doggy"). However, if the unfamiliar adjective was stressed ("GLIB doggy"), accurate localization of the named noun decreased (Thorpe and Fernald 2006), suggesting that prosody continues to play an important role in language perception. Swingley et al. (1999) showed that when 2-year-olds heard the word "dog," they more quickly looked at a picture of a dog when the alternate picture was of a tree, compared to a picture of a doll. In this situation, the picture of a doll (a known object) interfered with processing the object-label relation of "dog," presumably because the children required more lexical specificity before being able to make the correct choice.

As is evidenced in the aforementioned studies, much of the research in children older than 1 year of age has focused on lexical and grammatical development, with far less attention paid to continuing developments in speech perception *per se*, except with regard to atypical populations. Yet speech perception and discrimination abilities continue to evolve beyond the first year, as do children's strategies for integrating different sources of information (see, e.g., work by Nittrouer 1996 among others). Children's strategies for integrating different sources of information appear to shift as they place more weight on dynamically changing aspects of signals (and less weight on more static aspects) than do adult listeners (Robinson and Sloutsky 2004; Nittrouer and Lowenstein 2010). Many of these changes appear to be driven by differences in perceptual attention to components of the speech signal. Additional research with these older children, and particularly research examining the role of linguistic experience on perceptual abilities, is clearly warranted.

5 Perception of Speech in Challenging (Nonoptimal) Listening Conditions

The bulk of this chapter has summarized evidence on infants' abilities to attend to and extract information from the speech signal, and suggested that this ability is related to emerging knowledge of language structure. Moreover, growing interest in how early perceptual finesse predicts emerging language production importantly links these often separate research literatures. One final area of concern, however, is whether the kinds of experimental situations commonly employed in studies on infants' perception of speech adequately reflect the challenges faced by novice language learners. As discussed by Leibold (Chap. 5), infants and toddlers are often faced with multiple sound sources occurring simultaneously in modality-rich contexts, requiring then that speech be segregated from the background to make sense of the auditory signal. This final section of the chapter addresses different aspects of infants' speech perception under challenge: situations in which infants must focus and maintain attention to one of many sources of language-relevant inputs.

Several studies have reported that infants experience multitalker environments quite frequently (van de Weijer 1998), and thus their listening behavior in such settings may be a more realistic indication of their typical language exposure than is their listening behavior in a quiet laboratory setting. In fact, Trehub et al. (1981) found that infants' thresholds for detection of speech in noise were approximately 12 dB higher than those of adults, and Nozza et al. (1990, 1991) found that infants could distinguish phonemes in the presence of band-passed noise, but were more negatively affected than were adult listeners. These two studies, along with similar results for detecting tones in noise, suggest that infants' speech perception in noise is substantially poorer than that found in adults. Moreover, infants are not able to utilize top-down knowledge to "fill in" information masked by noise. When the source of noise is loud but brief (e.g., a car honking outside), such that it entirely masks a portion of the incoming signal, adult listeners "restore" the missing information, based on prior knowledge of the language (Warren 1970). Infants and toddlers do not show this same pattern (Newman 2006), and may be negatively affected by transient noises to a greater extent than adults.

More recent research has explored infants' speech perception in the presence of other types of noise, particularly natural sounds such as other people talking. In either word recognition tasks (e.g., Newman 2005), or segmentation tasks (Newman and Jusczyk 1996), the general conclusion is that infants' performance in the presence of background sounds is more compromised than that seen in adult listeners. Infants require substantially higher signal-to-noise ratios than adults in order to identify known words, especially during the first year of life (Newman 2005). Other studies suggest that infants' performance is also qualitatively different from adults' performance. For example, infants show better recognition when the distracter stream consists of multi-talker babble than of a single voice speaking, a pattern opposite that of adults (Newman 2009).

Another qualitative difference in perception between infants and adults is seen in the presence of auditory distracters that do not share acoustic space with speech (i.e., no acoustic masking). In one study, 6-month-old English-learning infants

showed no group-level discrimination of simple, native phonemes (e.g., *boo* vs. *goo*) in the presence of a high-frequency, natural distracter (Polka et al. 2008). However, the task itself had excluded factors (discussed previously) that most likely enhance infants' processing skills under challenge: the discrimination task involved a male, monotone voice, and a black-and-white checkerboard display. To extend this work, additional experiments by the first author (Panneton; in collaboration with Polka) have recently tested 6-month-olds' discrimination of the same easy phonemes, in the presence of the same distracters, but with various combinations of dynamic face and voice presentations. Preliminary results show that infants' discrimination in noise is improved when IDS is the style of vocal presentation. Similar studies have found that infants are better able to recognize familiar words in the presence of distracters when they can either see the face of the speaker (Hollich et al. 2005) or when the voice speaking is familiar to them (Barker and Newman 2004).

Collectively, these results reinforce the primary framework articulated earlier (Sect. 3) that the development of infants' perception of language is embedded in a typically rich, multimodal context, involving caretakers and others who adjust their style of communication to maximize infants' attention regulation, even if these adjustments flow naturally from emotional intentions. Presumably, infants who are better able to perceive speech in noise may be expected to have better language skills later on (another way in which perception in infancy can be related to emerging productive skills at later ages). According to the current view, individual differences in this domain most likely emerge in a variety of ways. One, the degree to which caretakers engage in sensitive, contingent, infant-directed styles of communication may enhance infants' resilience to the negative effects of noise and distracters on language processing. Second, variance in home-related ecologies (e.g., the overall amount of noise and distraction) may be related to infants' emerging adeptness at stream segregation. That is, living with more perceptual challenge may actually engender better performance at attending to a signal in noise. Clearly, future research extending these ideas to various populations of infants and young children is needed.

Lastly, infants and young children also face the ongoing challenge of processing speech emanating from speakers with whom they are not familiar (i.e., different vocal registers from their caretakers), who come from different natural groupings with distinct vocal signatures (e.g., males vs. females; adults vs. children), and who may be presenting their native language in unfamiliar accents or dialects. Initially, young infants fail to generalize familiar speech across genders, although they will generalize across talkers within a gender (Houston and Jusczyk 2000). Likewise, infants fail to generalize to speakers with novel accents (Schmale and Seidl 2009) and toddlers have difficulty recognizing words spoken in novel accents (Best et al. 2009). Thus, variability across talkers in gender or accent can pose difficulty for young infants, although with experience, such variation may actually afford infants the ability to normalize their representations of lexical forms such that they learn to ignore surface variability (Rost and McMurray 2010). Infants raised in bi-dialectal families may have greater opportunities to learn which acoustic variation is irrelevant than do infants raised in a single-dialect home. On the other hand, if accommodating dialectal differences requires additional cognitive resources from the child, it may

continue to pose problems, particularly in situations that are already difficult for the child (e.g., listening in noise) or which are near the limit of what the child is capable of doing at that stage of development.

6 Summary

This chapter began by placing the development of speech perception in a motivational framework, wherein infants' attention drives their language experience. Properties of caretakers' interactive styles encourage infants' attention, creating opportunities for learning about the native language. Language learning emerges from this dyadic context, with infants developing important skills from which they are able to bootstrap more advanced processing abilities. As a result, infants' ability to discriminate different languages, to segment the fluent speech stream, and to track statistical patterns in the input unfold, as attunement to native language properties increases. With greater experience, infants bring more advanced processing skills to the task of perceiving spoken language around them, progressively distributing their attention to essential aspects of the communicative environment.

Future work will contribute substantially to our understanding of language development if guided by awareness of important issues. First, extant research on children's speech perception has focused primarily on early skills, such as perceptual attunement and segmentation, and phonemic awareness in preschool children just before learning to read. New studies need to bridge infants' perceptual skills with children's developing production and comprehension skills. Second, future developmental studies need to address individual differences, rather than only group-level performance, in all domains of speech perception. The literature is dominated by studies on English or European language acquisition in monolingual families. More work is needed across disparate languages, and multilingual learning environments. Research is also needed on how speech perception skills differ in children at risk for clinical conditions. Moreover, not all learning takes place in the types of settings that mimic those in laboratories. Outside of the laboratory, infants face challenges such as the presence of noise, distracters, and signal variability. The degree to which infants accommodate perceptual challenge is not well understood, and may be an important source of individual differences in task performance. As perceptual challenge taxes cognitive resources, infants may fail to recognize or discriminate speech signals. As a result, infants' speech perception is situationally dependent as well as jointly influenced by individual processing abilities, motivation, and effective caretaking strategies (Werker and Curtin 2005).

This expanded focus opens up understanding of diverse pathways by which infants acquire their native language. Although nearly all children become successful language users, they do so via more than one trajectory. Examining individual differences will lead to more nuanced theories of how language acquisition builds on early perceptual skills and experiences. Addressing these questions will compel shifts in methodological approaches, including more graded responses

involving different functional systems, and multiple-point assessments. Developing methodologies that better enable collaborations, across laboratories as well as disciplines, will yield great benefits given inherent difficulties in securing infant samples with the size, diversity, and power to address new and important issues in this field.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: University of Edinburgh Press.
- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10, 48–53.
- Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Sciences*, 9, 92–98.
- Aslin, R. N., & Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: Promise, prospects, and challenges. *Journal of Biomedical Optics*, 10, 11009.
- Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, 94, B45–B53.
- Bernhardt, B. M., Kemp, N., & Werker, J. F. (2007). Early word-object associations and later language development. *First Language*, 27, 315–328.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345–360.
- Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, 20, 539–542.
- Bortfeld, H., Morgan, J. L., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16, 298–304.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2, 29–49.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46, 217–243.
- Boysson-Bardies, B. de, Sagart, L., & Durand, C. (1984). Discernable differences in the babbling of infants according to target language. *Journal of Child Language*, 11, 1–15.
- Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactics are useful for segmentation. *Cognition*, 61, 93–125.
- Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). Bilingual and monolingual infant phonetic development. *Applied Psycholinguistics*, 28, 455–474.
- Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2010). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, 16, 392–417.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221–268.
- Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, 1, 215–219.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics*, 31, 585–598.
- Cohen, L. B., Diehl, R. L., Oakes, L. M., & Loehlin, J. C. (1992). Infant perception of /aba/ versus /apa/: Building a quantitative model of infant categorical discrimination. *Developmental Psychology*, 28, 261–272.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, 52, 337–367.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174–1176.

- DeCasper A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*, 9, 133–150.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language*, 114, 53–65.
- Dietrich, D. E., Emrich, H. M., Waller, C., Wieringac, B. M., Johannes, S., & Münted, T. F. (2000). Emotion/cognition-coupling in word recognition memory of depressive patients: An event-related potential study. *Psychiatry Research*, 96, 15–29.
- Eimas, P. D., & Miller, J. L. (1980). Discrimination of information for manner of articulation. *Infant Behavior & Development*, 3, 367–375.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorrito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Fais, L., Kajikawa, S., Amano, S., & Werker, J. F. (2009). Infant discrimination of a morphologically relevant word-final contrast. *Infancy*, 14, 488–499.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46, 245–264.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behaviour and Development*, 8, 181–195.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27, 209–221.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mother's and father's speech to preverbal infants. *Journal of Child Language*, 16, 477–501.
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize words using partial phonetic information. *Child Development*, 72, 1003–1015.
- Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42, 98–116.
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67, 3192–3218.
- Friederici, A. D., & Oberecker, R. (2008). The development of syntactic brain correlates during the first years of life. In A. D. Friederici & G. Thierry (Eds.), *Trends in language acquisition research* (pp. 215–231). Amsterdam: John Benjamins.
- Gerken, L. A. (1996). Phonological and distributional information in syntax acquisition. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 411–425). Mahwah, NJ: Lawrence Erlbaum.
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268.
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19, 515–522.
- Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, 22, 703–726.
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? *Psychological Science*, 18, 254–260.
- Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). Developmental origins of voice processing in the human brain. *Neuron*, 65, 852–858.
- Hanulíková, A., McQueen, J. M., & Mitterer, H. (2010). Possible words and fixed stress in the segmentation of Slovak speech. *The Quarterly Journal of Experimental Psychology*, 63, 555–579.

- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581–616.
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preference for speech stimuli in Japanese infants. *Journal of Speech & Hearing Disorders*, 44, 1189–1200.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: An emergentist coalition model of word learning. *Monographs of the Society for Research in Child Development*, 65, 3 (Serial No. 262).
- Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, 76, 598–613.
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, 104, 353–366.
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1570–1582.
- Houston, D. M., Horn, D. L., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, 12, 119–145.
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13, 341–348.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548–568.
- Johnson, E. K., & Zamuner, T. S. (2010). Using infant and toddler testing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 73–94). Amsterdam: John Benjamins.
- Junge, C., Hagoort, P., Kooijman, V., & Cutler, A. (2010). Brain potentials for word segmentation at seven months predict later language development. In K. Franich, K. M. Iserman & L. L. Keil (Eds.), *Proceedings of the 34th annual Boston University Conference on Language Development* (Vol. 1, pp. 209–220). Somerville, MA: Cascadilla Press.
- Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1–23.
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645.
- Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal f0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development*, 67, 205–217.
- Kitamura, C., & Burnham, D. (1998). The infant's response to vocal affect in maternal speech. In C. Rovee-Collier (Ed.), *Advances in Infancy Research*, 12, 221–236.
- Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, 4, 85–110.
- Kitamura, C., & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy*, 14, 1–24.
- Klee, T., Pearce, K., & Carson, D. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 43, 821–833.
- Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24, 109–116.
- Kuhl, P. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10, 110–120.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, L. V., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684–686.
- Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the USA*, 100, 9096–9101.

- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 20, 215–225.
- Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6 F1–F10.
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near-infrared spectroscopy. *Neuroscience Biobehavioral Review*, 34, 269–284.
- Locke, J. (2001). First communion: The emergence of vocal relationships. *Social Development*, 10, 294–308.
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11, F9–F16.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10, 241–265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106, 1367–1381.
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception & Performance*, 27, 644–655.
- Mattys, S., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, 134, 477–500.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- McCardle, P., Colombo, J., & Freund, L. (2009). Measuring language in infancy. In Colombo, J., McCardle, P., & Freund, L. (Eds.), *Infant pathways to language: Methods, models, and research disorders* (pp. 1–12). New York: Psychology Press.
- McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6, 203–229.
- McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, 95, B15–B26.
- Mehler, J., Gervain, J., Endress, A. D., & Shukla, M. (2008). Mechanisms of language acquisition: Imaging and behavioral evidence. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed., pp. 325–336). Cambridge, MA: MIT Press.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Moon, C., Panneton-Cooper, R., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development*, 16, 495–500.
- Moore, D., Spence, M., & Katz, G. (1997). Six-month-olds' categorization of natural infant directed utterances. *Developmental Psychology*, 33, 980–989.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911–936.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility. *Psychological Science*, 15, 133–137.
- Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, 86, 157–199.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance*, 24, 756–766.
- Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43, 1–19.
- Nazzi, T., Paterson, S., & Karmiloff-Smith, A. (2003). Early word segmentation by infants and toddlers with Williams syndrome. *Infancy*, 4, 251–271.
- Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, 48, 279–298.

- Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, 41, 352–362.
- Newman, R. S. (2006). Perceptual restoration in toddlers. *Perception & Psychophysics*, 68, 625–642.
- Newman, R. S. (2009). Infant's listening in multitalker environments: Effect of the number of background talkers. *Attention, Perception, & Psychophysics*, 71, 822–836.
- Newman, R. S., & Hussain, I. (2006). Changes in infant preference for infant-directed speech in low and moderate noise by 4.5- to 13-month-olds. *Infancy*, 10, 61–76.
- Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception & Psychophysics*, 58, 1145–1156.
- Newman, R. S., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, 42, 643–655.
- Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech & Hearing Research*, 39, 278–297.
- Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America*, 127, 1624–1635.
- Nozza, R. J., Rossman, R. N. F., Bond, L. C., & Miller, S. L. (1990). Infant speech-sound discrimination in noise. *Journal of the Acoustical Society of America*, 87, 339–350.
- Nozza, R. J., Rossman, R. N. F., & Bond, L. C. (1991). Infant-adult differences in unmasked thresholds for the discrimination of consonant-vowel syllable pairs. *Audiology*, 30, 102–112.
- Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition*, 30, 583–593.
- Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow speech enhances younger, but not older infants' perception of vocal emotion. *Research in Human Development, Special Issue: The ecology of emotion in parenting relationships: Perspectives across the lifespan*, 3, 7–19.
- Panneton-Cooper, R. (1993). The effect of prosody on young infants' speech perception. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 137–167). Norwood, NJ: Ablex.
- Panneton-Cooper, R., & Aslin, R.N. (1990). Preference for infant-directed speech in the first months after birth. *Child Development*, 61, 1584–1595.
- Panneton-Cooper, R., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, 65, 1663–1677.
- Panneton-Cooper, R., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, 20, 477–488.
- Pegg, J., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, 15, 325–345.
- Peña, M., Makí, A., Kovačić D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the USA*, 100, 11702–11705.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Polka, L., & Bohn, O-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, doi:10.1016/j.wocn.2010.08.007.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421–435.
- Polka, L., Rvachew, S., & Molnar, M. (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy*, 13, 421–439.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 28, 349–351.
- Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, 75, 1387–1401.
- Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15, 680–635.

- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 58–108). Hoboken, NJ: John Wiley & Sons.
- Santesso, D. L., Schmidt, L. A., & Trainor, L. J. (2007). Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain and Cognition*, 65, 14–21.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, 46, 106–119.
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, 9, 235–248.
- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, 12, 583–601.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, 16, 197–201.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shifts in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874–887.
- Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science*, 9, 565–573.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3, 365–394.
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13, 410–420.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory & Language*, 49, 249–267.
- Soderstrom, M., Conwell, E., Feldman, N., & Morgan, J. L. (2009). Statistical learning in language acquisition: Beyond demonstrations, towards a theory. *Developmental Science*, 12, 409–411.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, 388, 381–382.
- Stern, D. (1985). *The interpersonal world of the infant*. New York: Basic Books.
- Sundara, M., & Scutellaro, A. (2010). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, doi:10.1016/j.wocn.2010.08.006.
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242.
- Swingle, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432–443.
- Swingle, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73–108.
- Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, 108, 850–855.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39, 706–716.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53–71.
- Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition*, 100, 389–433.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review*, 9, 335–340.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11, 188–195.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 97, 466–472.

- Trehub, S. E., Bull, D., & Schneider, B. A. (1981). Infants' detection of speech in noise. *Journal of Speech and Hearing Research*, 24, 202–206.
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, 120, 2285–2294.
- Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences of the USA*, 104, 13273–13278.
- van de Weijer, J. (1998). *Language input for word discovery*. Wageningen, the Netherlands: Ponsen & Looijen, BV.
- Vihman, M., Nakai, S., DePaolis, R. A., & Halle, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory & Language*, 50, 336–353.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10, 159–164.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316, 1159.
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning & Development*, 1, 197–234.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, 7, 49–63.
- Werker, J. F., Polka, L., & Pegg, J. E. (1997). The conditioned head turn procedure as a method for assessing infant speech perception. *Early Development and Parenting*, 6, 171–178.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34, 1289–1309.
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103, 147–162.
- White, L., Melborn, J. F., & Mattys, S. (2010). Segmentation by lexical subtraction in Hungarian speakers of second-language English. *Quarterly Journal of Experimental Psychology*, 63, 544–554.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Zangl, R., Klarman, L., Thal, D. J., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, 6, 179–208.

Chapter 8

Development of Pitch and Music Perception

Laurel J. Trainor and Andrea Unrau

1 Introduction

Music, like language, is found in all known human societies, past and present, and although music has some precursors that can be found in other species, it is a defining characteristic of the human species (Wallin et al. 2000). Although the perception of music has multisensory aspects (Thompson et al. 2005; Phillips-Silver and Trainor 2007), it is based in sound. The vibrations that give rise to music can be created by many different means, including striking percussion instruments, blowing resonating air columns in wind instruments, plucking and bowing strings under tension, and vibrating the vocal chords during singing. The boundary between sounds that are perceived as music and sounds that are not is fuzzy; indeed a definition of music remains elusive. However, the vast majority of music in the world involves spectral and temporal organizations that can readily be processed by the nervous system and for which specialized brain processing develops.

Although the evolutionary adaptive value of music remains controversial (McDermott and Hauser 2005; Trainor 2008a), even in modern societies people spend large amounts of time and money on music (Huron 2003), suggesting that it serves important functions. From a neurological perspective, music is an interesting stimulus as it activates many cortical areas, including sensorimotor systems (e.g., Grahn and Brett 2007; Zatorre et al. 2007), memory and cognitive systems (Koelsch and Siebel 2005; Janata 2009), and social/emotional areas (e.g., Blood and Zatorre 2001). Music appears to have a special relation with movement, as even passive listening activates motor areas (Grahn and Brett 2007), and the way people move can affect how the rhythm of a musical stimulus is interpreted (Phillips-Silver and

L.J. Trainor (✉) • A. Unrau
Department of Psychology, Neuroscience & Behaviour, McMaster University,
Hamilton, ON, Canada
e-mail: LJT@mcmaster.ca; unrauaj@mcmaster.ca

Trainor 2005, 2007). From a cognitive perspective, there is evidence that superior musical skills are accompanied by higher IQ scores (Schellenberg 2005), better mathematical ability (e.g., Vaughn 2000), and enhanced language development (e.g., Anvari et al. 2002; Magne et al. 2006; Moreno and Besson 2006). Perhaps most intriguing are recent reports that music enhances social relationships. Children who have sung together while playing show more prosocial behavior than children who have simply played together (Kirschner and Tomasello 2010). Even in infancy, music is used for social/emotional communication (Trehub 2009). Singing to infants is universal across cultures (Trehub and Trainor 1998), and infants prefer the loving tone of voice and friendly higher pitch of infant-directed compared to non-infant-directed singing (Trainor 1996; Trainor and Zacharias 1998). Indeed, caregivers use singing to achieve various parenting goals. The soft, slow, airy, low-pitched characteristics of lullabies put infants to sleep and the fast, high-pitched, rhythmic characteristics of play songs rouse infants and encourage them to engage in social interaction (Trainor et al. 1997).

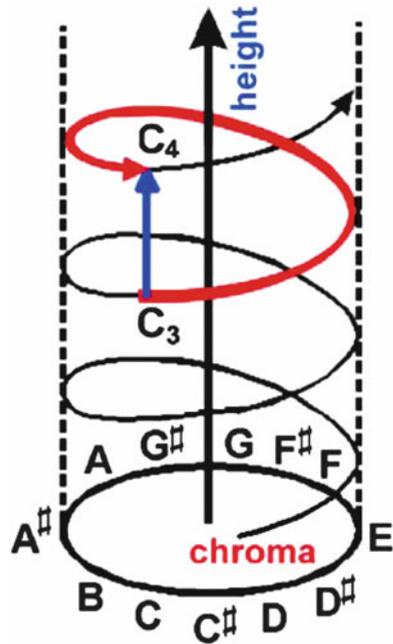
To understand why perceiving and producing music are such powerful human experiences, it is necessary to consider how the perception of music develops. Musical structure consists of two basic aspects, a pitch aspect and a rhythmic aspect. The present chapter focuses on the development of musical pitch processing. It also focuses primarily on Western musical structure, because this is the only musical system for which there is substantial developmental data. The following sections examine behavioral and neural development for musical pitch processing and the effects of experience on development.

2 Musical Pitch Structure: Universal and System-Specific Aspects

Before considering the development of sensitivity to musical pitch in subsequent sections, here the various aspects of musical pitch structure are outlined. Which aspects are universal across musical systems and which aspects are specific to Western musical structure are also considered. A general prediction that can be made is that universal musical elements likely reflect aspects of auditory processing that capitalize on features of the auditory system; therefore, these elements could be fairly easy to learn and might develop relatively early. On the other hand, features that are specific to one musical system might reflect more mature ways of processing and, therefore, require more experience and develop later.

Musical pitch is complex and hierarchical in nature (Shepard 1964; Krumhansl 1990). At the most basic level, individual musical tones have a particular *pitch height*, ordered from low to high. This aspect of pitch processing is universal and domain general, applying, for example, to speech as well as music. A second level of the pitch hierarchy arises because tones with fundamental frequencies an octave apart, a doubling of frequency, are perceived as similar. Indeed, in Western musical notation, notes an octave apart are given the same note name (e.g., C, D, E) and are

Fig. 8.1 Helical representation of pitch similarity. Pitch height is represented in the vertical dimension and chroma around the circle. The distance between two points represents their perceived similarity, accounting for both the similarity of adjacent tones and tones separated by octaves [Adapted from Shepard (1965). Reprinted by permission]



referred to as *chroma*. When listeners are asked to rate the similarity of pairs of tones, they take into account both pitch height and pitch chroma, and perceived similarity can be represented geometrically by a helix (Shepard 1982) (Fig. 8.1). Although the particular chroma used differ across musical systems, the height and chroma dimensions of musical pitch appear to be universal.

The perceptual similarity of notes an octave apart is undoubtedly related to another fundamental and universal aspect of pitch perception, that of the continuum between *consonance* and *dissonance*. Most simply defined, two notes played simultaneously are said to be consonant when they sound pleasant or smooth together whereas they are said to be dissonant when they sound rough or unpleasant. The octave is the most consonant musical interval. Consonance and dissonance contribute to people's emotional experiences of music because dissonance increases musical tension and consonance releases this tension.

Pitch is encoded in the nervous system via two mechanisms, a temporal code corresponding to the repetition rate or frequency of the incoming sound, and a spatial code whereby different frequencies are represented across neural tissue in an orderly fashion to produce tonotopic maps. Clear tonotopic maps are found in all auditory subcortical nuclei as well as in primary auditory cortex. Despite this absolute frequency organization, most people are rather poor at remembering the exact or *absolute pitch* of individual tones. On the other hand, people are generally quite good at remembering pitch intervals—the distance between two pitches—or *relative pitch*. For example, people recognize a familiar tune, such as “Happy Birthday”,

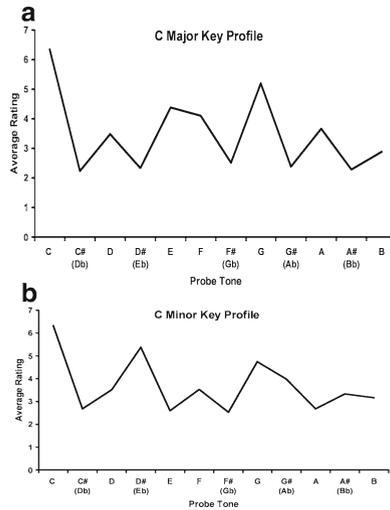


Fig. 8.2 Probe tone ratings for the keys of (a) C major and (b) C minor. Experienced musicians heard a key context (either a tonic triad or one of three chord cadences) and were then asked to rate how well a probe tone (taken from all 12 notes found within an octave in Western music) fit that context. For both major and minor contexts, tonic tones were always rated highest, followed by the other notes within the scale. Non-scale tones were always rated lowest [Adapted from Krumhansl and Kessler (1982) with permission of the American Psychological Association]

regardless of the absolute pitch of the starting note, as long as the intervals or pitch distances between the notes of the song are correct. This phenomenon, known as *transpositional invariance*, indicates that music is largely encoded in long-term memory in terms of relative rather than absolute pitch. The dominance of relative pitch encoding and transpositional invariance are likely universal across musical systems, although the prevalence of the use of absolute pitch encoding changes with culture and is likely influenced by experience (e.g., Deutsch et al. 2004).

One very interesting aspect of musical pitch structure is that music is generally not composed using continuous pitch but, rather, using a small number of *discrete pitch categories*, subdivisions of the octave. The collection of pitch categories forms a musical *scale*. The use of scales is near universal, but the particular scales used vary considerably from one musical system to another. In the most common Western scales, each note of the scale relates to the others in a unique way and takes on a specific function; the idea that different notes of the scale have different degrees of stability or importance is referred to as the *tonal hierarchy* (Fig. 8.2).

Because of transpositional invariance, a scale can begin on any note and it retains its identity as long as it contains the correct structure of pitch distances or intervals between notes of the scale. In Western music, 12 major scales can be created starting on each of the notes formed by subdividing the octave into 12 equal intervals (on a log scale) called semitones. Each one of these scales is called a musical *key*, and Western musical compositions are typically in one key at any particular time. Each of the 12 keys is most closely related to two other major keys with which it shares

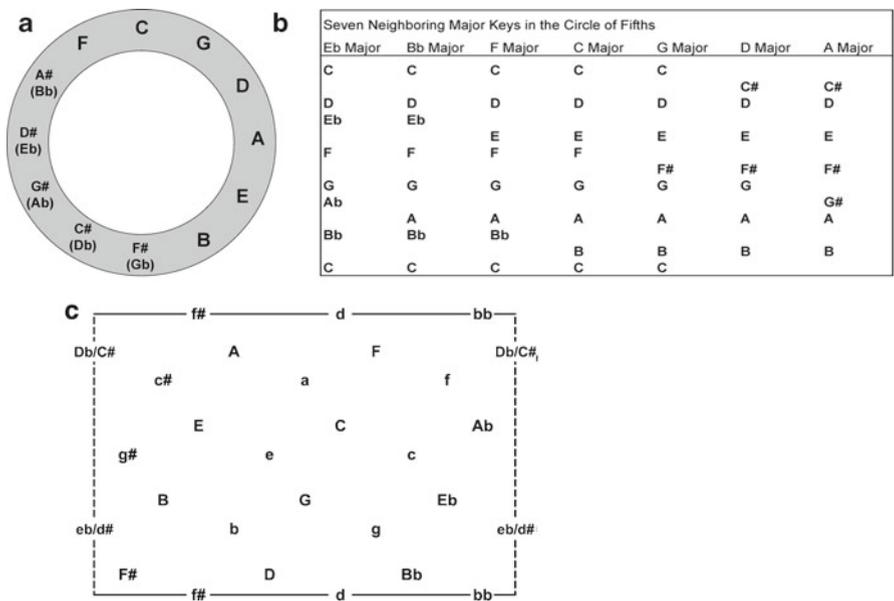


Fig. 8.3 Representations of key relatedness. (a) Western major keys can be represented on a circle in which adjacent keys are most closely related. (b) Adjacent keys are separated by a perfect fifth and share all but one scale note. C major, for example, shares all scale notes except F natural with G major, and all but B natural with F major. (c) The rectangular map is a two-dimensional representation of the relatedness between all 24 Western major (*uppercase*) and minor (*lowercase*) keys. The *top* and *bottom* of the map are identical, as are the *left* and *right* edges, so this is a two-dimensional representation of a three-dimensional *torus* [Adapted from Krumhansl and Kessler (1982) with permission of the American Psychological Association]

6 of its 7 notes. The 12 keys can be arranged in a circle such that the farther away two keys are on the circle, the fewer notes they share and the less related they are (Fig. 8.3a, b). Transpositions from one key to another are most likely to occur between related keys. Western musical structure also includes a minor key associated with each major key. The major and natural minor keys contain the same notes, but the tonic is shifted down two notes in the minor relative to the major version. The major and minor keys can be represented in a map of tonal pitch space in which the distance between keys in the space corresponds to how closely they are related (Fig. 8.3c). Physiologically, the representation of musical keys appears to involve complex brain networks, including frontal brain regions (Janata et al. 2002).

In addition to the sequential pitch structure of melodies, Western music has prominent *harmonic structure*. Indeed, each note of a melody can be associated with a chord as when chord-based accompaniment is provided. Chords are typically built on the seven notes of the scale. Harmonic structure has a kind of loose syntax whereby certain chords are expected to follow other chords in sequence. Violations of harmonic expectations are frequently used in musical composition, and can give

rise to emotional experiences in listeners (Meyer 1956; Huron 2006). Even musically untrained Western listeners are quite sensitive to harmonic expectations, showing smaller event-related potential (ERP) brain responses to expected chords compared to unexpected chords (Koelsch et al. 2000).

In summary, although musical pitch is based on domain-general mechanisms for pitch processing, there are many layers of pitch structure in music. Some of these layers are very common across musical systems, such as the dimensions of pitch height and chroma, consonance and dissonance, transpositional invariance, the use of discrete pitch categories, scales with different sizes of intervals, and tonal pitch spaces. However, different musical systems use different musical scales and tonal pitch spaces, and the use of complex harmony appears to be largely limited to Western musical structure. In the next section the development of these aspects of musical pitch is explored.

3 Development of Musical Pitch Processing

3.1 *Basic Pitch Perception*

Sounds with energy at only one frequency are referred to as pure or sine tones. However, sounds with pitch are typically complex in that they have energy at a fundamental frequency, corresponding to the perceived pitch, and at harmonics at integer multiples of that frequency. Normally, the frequency components are fused into a single percept. All components contribute to the perceived pitch, as the pitch does not change when the fundamental frequency is removed.

As outlined by Buss, Hall, and Grose (Chap. 4), young infants are able to discriminate tones of different frequencies although their performance is poorer than that of adults. However, discrimination is good enough to support the perception of musical pitch. A related question concerns whether infants and adults use similar cortical mechanisms for discriminating tones. This question can be addressed by measuring event-related potentials (ERPs) derived from electroencephalogram (EEG) recordings. The general immaturity of auditory cortex in young infants is seen in that the frequency of the dominant oscillatory rhythm from auditory regions increases from 4 to 5 Hz at 4 months to 6–7 Hz at 12 months to around 10 Hz in adults (Fujioka et al. 2011). In adults, presentation of a tone triggers a series of responses over a few hundred milliseconds (ms) that represent the stages of sound processing (Fig. 8.4), beginning with auditory brainstem responses (within the first 15 ms after sound onset), followed by middle latency responses from primary auditory cortex (<50 ms after sound onset), followed by sensory, perceptual and cognitive cortical responses, including P1 (the first large frontally positive response ~50 ms), N1 (the first large frontally negative response ~100 ms), P2 (~170 ms), N2 (depending on attention, ~250 ms) and P3 (relating to conscious performance of a task, at various latencies depending on the task). Activity in auditory cortex generally gives rise to a dipolar pattern on the surface of the head, with a negativity at

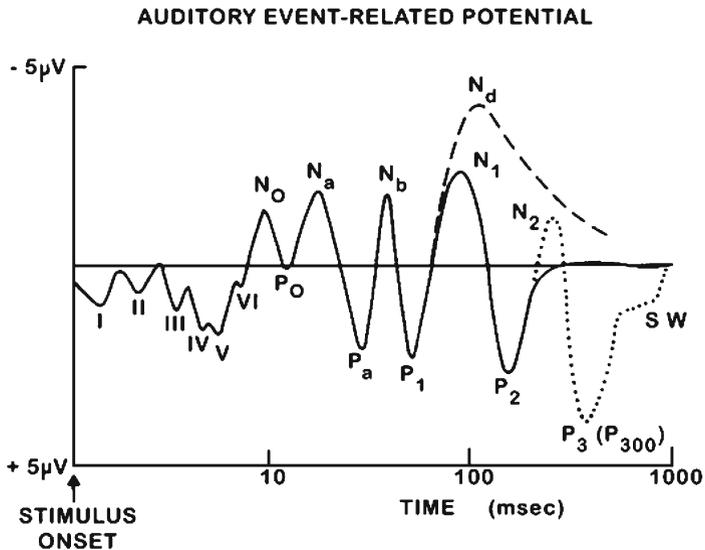
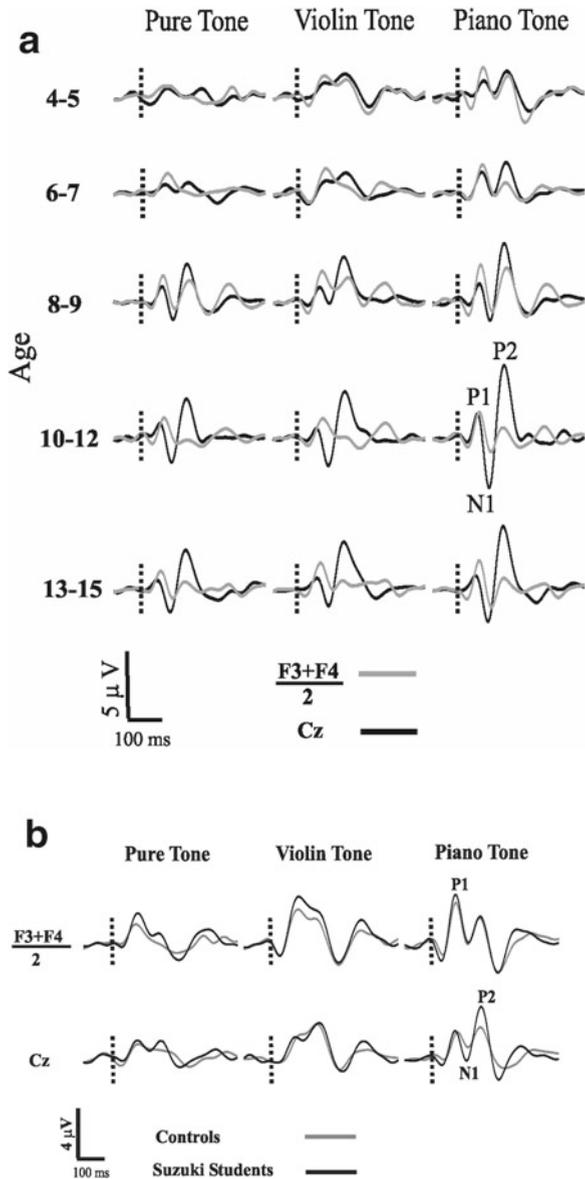


Fig. 8.4 Stylized representation of the auditory event-related potential (*ERP*) from a frontal site on the surface of the head. The auditory *ERP* shows the stages of processing in the brain's response to a sound and is derived from electroencephalogram (*EEG*) recordings. The first few peaks (I–VI) represent auditory brain stem responses; peaks N0, P0, Na, Pa, Nb, and P1 represent the middle latency responses from primary auditory cortex; N1 and P2 represent processing from secondary auditory areas. N2, Nd, and P3 are nonobligatory responses reflecting attention, memory and decision processes [From Trainor (2008b). Reprinted with permission of Cambridge University Press]

frontal sites accompanied by a positivity at posterior sites, or vice versa. Interestingly, these components take a very long time to mature. Specifically, the N1 and P2 components are so small as to be difficult to measure before about 4 years of age. They decrease in latency and increase in amplitude until about 10–12 years of age, and decrease thereafter, reaching stable adult levels in the late teenage years (Ponton et al. 2000; Shahin et al. 2004) (Fig. 8.5a). This protracted development suggests that these components represent feedback connections from other brain regions that provide top-down signals to auditory cortex. After early hearing deprivation, these components do not develop normally (Eggermont and Moore, Chap. 2).

Despite the late emergence and slow maturation of the obligatory responses to sound described in the previous paragraph, ERP components reflecting the brain's response to occasional changes in an ongoing stream of sound emerge much earlier (Trainor 2008b). Indeed, the main task of the perceptual system can be seen to be one of predicting the future based on the past, and performing extra processing to update representations of the state of the world when a prediction is incorrect (Huron 2006; Trainor and Zatorre 2009). In adults, virtually any occasional detectable change (e.g., in pitch, intensity, timbre, pattern) in an ongoing sequence of sounds gives rise to an additional negativity, termed a mismatch negativity (MMN), peaking at frontal sites on the scalp between 120 and 250 ms after stimulus onset depending on the complexity

Fig. 8.5 Development of auditory event-related potential (ERP) components. **(a)** The protracted development of the N1 and P2 ERP components is shown in response to pure tones, violin tones and piano tones. The *dotted lines* represent sound onset. Components increase in amplitude up to 10–12 years of age, and then gradually decrease to stable adult levels over the teenage years. **(b)** 4- and 5-year-old children engaged in musical training show enhanced N1 and P2 responses compared to control participants not training musically. F3 and F4 are *left* and *right* frontal scalp regions, respectively, and Cz is a central scalp region [From Shahin et al. (2004). Adapted with permission of Wolters Kluwer/Lippincott, Williams & Wilkins]



of the change. The ERP responses of young infants to single tones are dominated by frontally positive slow waves, components that disappear with increasing age. At 2 months of age, occasional changes in pitch primarily elicit an increase in the amplitude of this frontally positive slow wave (He et al. 2007, 2009). However, by 3 months an MMN with adult-like morphology is evident, which decreases in latency and increases in amplitude with increasing age (Fig. 8.6). Thus, young infants are

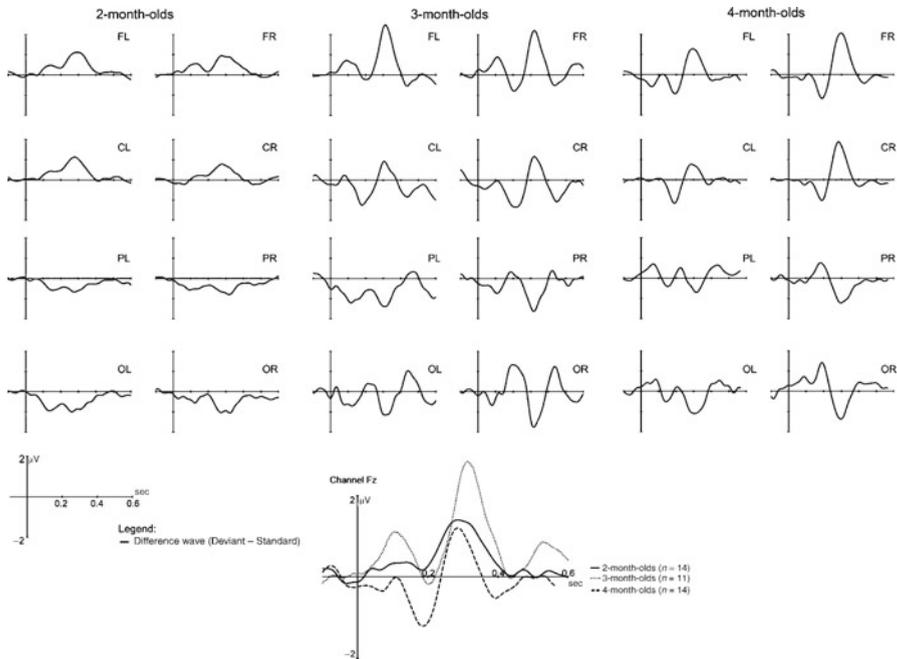


Fig. 8.6 Event-related potential (ERP) responses in early infancy to occasional pitch changes. On standard trials (80%) a C5 piano tone (523 Hz) was presented and on deviant trials (20%) an F#5 piano tone (740 Hz) was presented. Shown are the ERP difference (deviant-standard) waves at left (*L*) and right (*R*), frontal (*F*), central (*C*), parietal (*P*), and occipital (*O*) scalp regions. The y-axis represents tone onset. The ERPs reverse in polarity from front to back of the head, consistent with a neural generator in auditory cortex, which is located between these sites. In response to the change in pitch, 2-month-olds show only an increase in a frontally positive slow wave, whereas 3- and 4-month-olds show a faster, frontally negative wave much like the adult mismatch negativity (MMN), followed by a positivity much like the adult P3a. [From He et al. (2007). Reprinted with permission of MIT Press]

able to discriminate tones with different fundamental frequencies, but an adult-like cortical mechanism for this takes a few months to emerge.

When processing complex tones, the auditory system breaks incoming sounds into their frequency components, and much subcortical processing takes place within frequency-specific channels. Such analysis of the frequency content of sounds is crucial for separating simultaneous sound sources with overlapping frequency content and for the perception of timbre (Leibold, Chap. 5). Evidence from monkeys suggests that the harmonics of a complex tone are not integrated into a single pitch representation subcortically or in primary auditory cortex (Fishman et al. 1998), but that pitch is first extracted in a region adjacent to primary auditory cortex (Bendor and Wang 2005). Studies in humans examining functional magnetic resonance imaging (fMRI) responses (Patterson et al. 2002; Penagos et al. 2004) and effects of lesions (Schönwiesner and Zatorre 2008) indicate a similar organization.

Given the immaturity of auditory cortex during the first months after birth, it might be expected that very young infants can discriminate pure tones, but cannot initially integrate harmonics into a single pitch percept. Complex pitch perception can be studied with stimuli in which the fundamental frequency has been removed, because if the fundamental is present, there is no way of determining whether the subject is integrating all of the harmonics into a pitch percept or attending only to the fundamental frequency. If removal of the fundamental does not change the perception of the pitch, then it is clear that the subject must have integrated the harmonics in order to perceive the pitch. Clarkson and Clifton (1985) presented 7-month-old infants with different tones containing various harmonics but all missing the fundamental. They used a *conditioned head turn procedure* in which the set of missing-fundamental tones was presented in random order from a loudspeaker to the side of each infant. Occasionally a tone with a different pitch was presented, and if infants turned their heads at least 45° toward the sound source when this occurred, they were rewarded with dancing toys for a couple of seconds. Head turns at other times were not rewarded, and discrimination was determined by comparing the proportion of head turns when there was a change in pitch to when there was not. Clarkson and Clifton found that the infants were able to ignore the timbre changes and respond to changes in the pitch of the missing fundamental. At the same time, the infants, unlike adults, were unable to do the task if only high harmonics were present (Clarkson and Rogers 1995), indicating that basic pitch perception is not yet adult-like at 7 months.

He and Trainor (2009) used ERPs to investigate the emergence of pitch perception. They presented standard tone pairs in which the fundamental and different harmonics were present on each trial. In each standard tone pair, the frequency of all harmonics, as well as the overall pitch, increased from the first to the second tone, although the pitch of the first tone and the increase in frequency varied from pair to pair. Occasionally He and Trainor presented a deviant tone pair in which the frequency components of the second tone lined up to produce the pitch of a missing fundamental frequency that was lower than the pitch of the first tone (Fig. 8.7a). Note that in deviant tone pairs the frequency of each component increased from the first to the second tone just as in the standard stimuli, so that if the pitch of the missing fundamental was not perceived, deviant stimuli would generate similar ERP responses as standard stimuli. He and Trainor found that adults, 7-month-olds and 4-month-olds responded to the pitch reversal in the deviant pairs, but found no evidence that 3-month-olds did (Fig. 8.7b). Thus it appears that the ability to integrate harmonics into a pitch percept emerges between 3 and 4 months of age.

In sum, although frequency discrimination is good enough in early infancy to support musical pitch processing, auditory cortex remains immature for a protracted period. ERP components related to communication between auditory areas and other brain regions do not mature until the late teenage years. The ability to integrate the harmonics of a complex tone into a unified pitch representation relies on brain regions beyond primary auditory cortex and may not emerge until several months after birth.

3.2 Consonance and Dissonance

The consonance/dissonance continuum is one of the fundamental organizing principles of musical pitch. In adults, increases and decreases in consonance and dissonance give rise to increases and decreases in perceived tension, which in turn give rise to the ebb and flow of perceived emotion in music (Smith and Cuddy 2003). One explanation of the perceptual difference between consonance and dissonance involves the structure of the inner ear. Tones whose fundamental frequencies stand in simple integer ratios, such as the octave or perfect eighth (2:1 ratio) or the perfect fifth (3:2 ratio), tend to sound consonant whereas tones whose fundamental frequencies stand in more complex integer ratios, such as the major seventh (15:8) or tritone (45:32), tend to sound dissonant. Most of the lower, most salient harmonics of consonant tone pairs are either the same or are separated by more than a critical bandwidth (about one-fourth of an octave for much of the frequency range). In contrast, highly dissonant tone pairs have many harmonics between them that are less than a critical bandwidth apart. When two frequency components across simultaneous complex tones (whether fundamentals or harmonics)

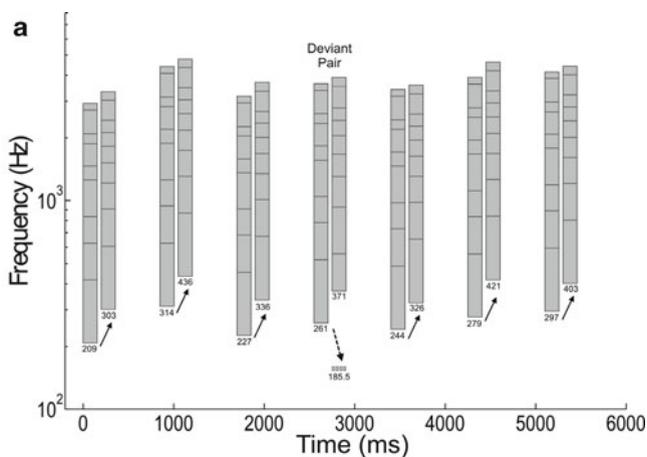


Fig. 8.7 Processing the missing fundamental in infancy. **(a)** The stimuli were complex tones (*gray bars*) created using 10 harmonics (*black lines* within the *gray bars*) from the first 15 harmonics. All trials consisted of a tone pair in which every harmonic increased from the first to the second tone of the pair. For standard pairs (80%), the pitch corresponded to the lowest harmonic (fundamental), and the pitch increased from the first to the second tone. However, in the case of deviant pairs (20%) the harmonics of the second tone in the pair were all integer multiples of a missing fundamental that was lower than the frequency of the lowest harmonic in the first tone. Therefore, if the pitch of the missing fundamental was perceived, the pitch rose from the first to the second tone in the standard pairs, but the pitch fell in the case of deviant pairs. However, if the harmonics were not integrated into a pitch percept, standard and deviant pairs would not be distinguishable as all individual harmonics rose from the first to the second tone for both trial types.

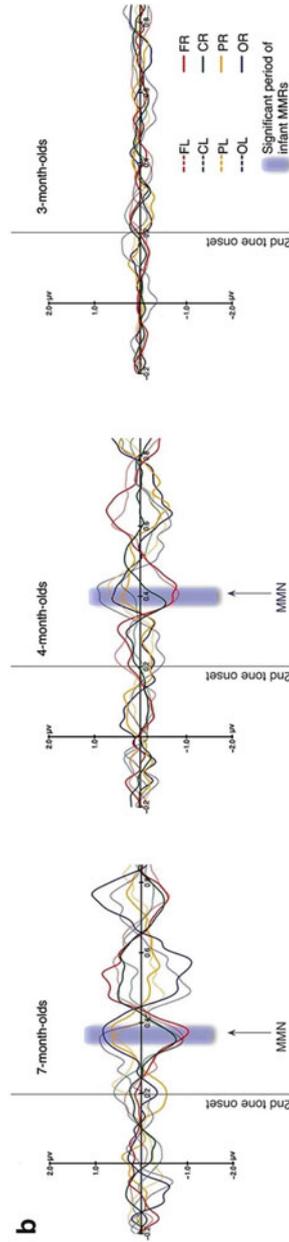


Fig. 8.7 (continued) (b) Event-related potential (ERP) difference waves (deviant – standard trials) at left (*L*) and right (*R*), frontal (*F*), central (*C*), parietal (*P*), and occipital (*O*) scalp regions. The mismatch negativity (MMN) component (highlighted by the *gray bars*) can be seen at 7 and 4 months of age by the frontal negativity simultaneous with a posterior positivity in the difference waves. Thus, 7-month-old and 4-month-old infants, but not 3-month-old infants, showed evidence of cortical response to the pitch of the missing fundamental [From He and Trainor (2009). Reprinted with permission of the Society for Neuroscience]

are less than a critical bandwidth apart they cannot be represented independently on the basilar membrane in the inner ear, and the complex patterns of activity they elicit give rise to the sensation of beating or roughness. It has been proposed that it is the roughness and beating of such stimuli that give rise to the sensation of dissonance (Plomp and Levelt 1965).

There is also evidence that the consonance/dissonance distinction arises from the distinct temporal firing patterns that consonant vs. dissonant intervals set up in the auditory nerve fibres (Tramo et al. 2001). Indeed a recent study suggests that it is the harmonicity (i.e., whether harmonics stand at integer multiples of a fundamental) rather than roughness that plays the greatest role in determining perception of consonance versus dissonance (McDermott et al. 2010). Regardless of whether the perception of consonance and dissonance is the result of complex vibration patterns on the basilar membrane or temporal firing patterns in the auditory nerve, its origin is relatively peripheral in the nervous system in both cases. Consistent with this, the perception of consonance and dissonance appears to be quite similar across cultures for isolated tones, although effects of the surrounding musical context on consonance perception may differ substantially across cultures.

If the distinction between consonance and dissonance is peripheral in origin, sensitivity might be expected to emerge early in development. Indeed, infants, like adults, are better able to process consonant than dissonant intervals. When highly consonant intervals, each consisting of two simultaneous tones, are presented in random order, infants readily detect when a dissonant tone is inserted, but they are not able to detect a consonant interval inserted into a set of dissonant intervals (Trainor 1997). Infants are also better at detecting a semitone change from a perfect fifth interval that creates a dissonant interval compared to a larger whole tone change that creates another consonant interval (Schellenberg and Trainor 1996). Finally, in melodic contexts, infants can more easily detect changes in melodies with prominent consonant intervals compared to melodies with prominent dissonant intervals (Trainor and Trehub 1993). Together, these studies indicate that the consonance/dissonance continuum is perceptually salient to infants and has consequences for the processing of tonal information.

It is also of interest that in addition to a processing advantage for consonance, infants prefer consonance over dissonance. Using looking time preference procedures in which infants determine how long they listen to the sound presented on each trial by how long they look at a visual display—the sound for each trial ends when the infants look away from the visual display—it has been revealed that infants as young as 2 months prefer to listen to consonant over dissonant intervals in isolated contexts or in the context of simple musical pieces (Trainor and Heinmiller 1998; Zentner and Kagan 1998; Trainor et al. 2002b). Indeed, this preference is found very early and without prenatal speech exposure in hearing newborns of deaf parents (Masataka 2006).

The origin of the preference for consonance over dissonance is not known. It could possibly be innate, or it could arise from experience with the harmonics of complex tones, in which the lower, most salient harmonics form consonant intervals (Terhardt 1974). It might also be related to ease of processing. Simultaneous tones that form dissonant intervals have harmonics that are close in pitch but not identical.

Such harmonics are processed within a critical band and therefore are not represented completely independently on the basilar membrane making it difficult to determine the exact frequencies present. Whatever the case, it is clear that sensitivity to the consonance/dissonance continuum arises very early in development.

Much of musical pitch structure relates back to the consonance/dissonance continuum. *Octave equivalence*, that tones an octave apart sound similar, is a near-universal feature of musical systems (Burns 1999) and likely arises because tones an octave apart share most of their harmonics, and because the octave is the most consonant interval. When people sing together and their voices are in different pitch ranges, they normally sing octaves apart. Like consonance and dissonance, infants are sensitive to octave relationships (Demany and Armand 1984). Octave equivalence, in turn, enables the emergence of musical scales, which divide the octave into discrete intervals. Musical scale intervals are also influenced by the consonance/dissonance continuum as the more prominent intervals in musical scales are also the most consonant (e.g., the interval between the two most stable notes of the Western major scale is a perfect fifth or 3:2 ratio).

In sum, sensitivity to consonance and dissonance emerges early, as does a preference for consonance. Because of its relation to octave equivalence and musical scale structures, it is possible that the consonance/dissonance continuum is the fundamental organizing principle for musical pitch structures.

3.3 *Transpositional Invariance*

Transpositional invariance is not specific to musical pitch structure in that people readily recognize, for example, speech sounds, whether spoken in a high or low pitch range. But transpositional invariance is crucial for musical pitch structure, first, because it enables recognition of melodies across a wide variety of contexts and, second, because melodies often contain short pitch patterns that are repeated at different pitch levels, which give melodies a sense of cohesion. Transpositional invariance requires ignoring the absolute pitch information and attending to the relative pitch distances between tones. Interestingly, young infants have transpositional invariance in that they treat melodies transposed to different starting notes as similar, readily detecting occasional differences in relative pitch structure despite the fact that the transposed versions being compared have no actual notes in common, only relative pitch distances (e.g., Trehub et al. 1984; Trainor and Trehub 1992; Chang and Trehub 1977). Not only can adults recognize melodies in transposition, but it is the normal mode of processing for most adults. There is evidence that infants (Saffran and Griepentrog 2001; Volkova et al. 2006) and adults (Schellenberg and Trehub 2003) can remember absolute pitch under some circumstances, but for both age groups, relative pitch and transpositional invariance appear to be the normal mode of processing. Although access to absolute pitch is maintained in working memory, for most people, absolute pitch information fades rapidly. Given a target tone to remember, the presentation of interference tones with random pitch dramatically reduces most adults' ability to identify the pitch of the target tone (Ross et al. 2004). Similarly,

infants' ability to compare the pitch of two tones decreases with an increasing number of random pitch interference tones (Plantinga and Trainor 2008).

Infants' long-term memory representations for melodic pitch are also primarily in terms of relative pitch. Plantinga and Trainor (2005) tested this by playing a few repetitions of a melody daily for infants over the course of a week and then examining their listening preferences. Half of the infants were familiarized with one melody whereas the other half heard a second melody. After a 1-day delay between last familiarization and testing, infants looked longer on trials presenting the unfamiliar melody compared to the familiar melody, indicating that they remembered the melodies. Of most interest here, in a second condition, both melodies were transposed up or down by either a perfect fifth or a tritone. Under transposition, every note of each melody is different, making recognition dependent on a relative pitch representation. Preferences for the novel melody under transposition were as strong as when familiarization and test melodies were at the familiarized pitch, indicating robust relative pitch representations. Furthermore, when the familiar melody was presented in two versions during the preference test, one at the pitch heard during familiarization and one transposed, infants showed no preference, again indicating that either they do not have a long-term representation for the absolute pitch or it is not salient to them.

ERP data reveal that transpositionally invariant representations of melodies can be found in auditory cortex at 6 months of age. When a four-note melody was presented in transposition, occasional wrong notes in the final position elicited brain responses, indicating that infants detected the changes (Tew et al. 2009). Interestingly, although occasional changes in the pitch of a single repeated tone elicited MMN responses with an adult-like morphology by 3 months of age as discussed above, for melodies, the mismatch response remained immature even at 6 months. Specifically, instead of a frontally negative/posteriorly positive response to the deviant tones as seen in adults, when the melodic/transpositional context required relative pitch processing, 6-month-olds showed an increase in the frontally-positive slow wave in response to the deviant tones. This response is similar to that shown by younger infants in response to simple pitch changes (see Sect. 3.1). Thus, although infants process relative pitch early in development, the cortical mechanisms and sound representations remain immature for some time. It is not yet known when they mature.

In sum, as early as has been tested, infants show robust processing of relative pitch in both short-term and long-term memory representations. This ability is domain-general to an extent, being important for detecting invariance in speech and various environmental sounds. Its early emergence is likely an important prerequisite for musical acquisition.

3.4 Learning Musical Scales and the Tonal Hierarchy

Music is composed from sets of discrete notes or scales. The most common scale in Western music is the major scale. It contains two interval sizes between adjacent notes of the scale, the tone (1/6th of an octave) and the semitone (1/12th of an

octave), such that it divides the octave into an ordered set of seven intervals: tone, tone, semitone, tone, tone, tone, semitone. Although the use of scales is near universal, the particular scales used vary across musical systems. Therefore, it is necessary to learn the particular musical scales of one's culture, just as it is necessary to learn the language of one's culture. And just as children learn the language of their culture through exposure but no formal training, children acquire implicit knowledge about the scale structure of their culture through mere exposure. Adults without formal musical training and no explicit knowledge of musical nomenclature and structure have considerable implicit knowledge about the musical system to which they have been exposed (e.g., Trainor and Trehub 1994; Bigand and Poulin-Charronnat 2006; Bischoff Renninger et al. 2006; Tillmann et al. 2006).

By at least as young as 2 months of age, infants can discriminate unfamiliar melodies (Plantinga and Trainor 2009) and by 6 months infants can remember melodies for weeks (Saffran et al. 2000; Trainor et al. 2004; Ilari and Polka 2006). On the other hand, it takes infants exposed to Western music some time to learn key membership, or which notes belong in a key. Trainor and Trehub (1992) presented adults and 8-month-old infants with a ten-note melody and measured their ability to detect occasional note changes. Adults were much better at detecting changes to notes that did not belong to the key of the melody in comparison to changed notes that remained within the key. On the other hand, infants were equally good at detecting both change types, indicating that they had not yet acquired knowledge of key membership for the Western major scale. Lynch et al. (1990) examined cross-cultural processing. They found that Western adults were much better at detecting changed notes in melodies based on the Western major scale compared to melodies based on an unfamiliar Balanese scale. Infants, on the other hand, performed equally well with melodies in the two scales. The exact age at which children acquire musical system-specific knowledge remains unknown, and likely depends on the amount and type of music experience they have. Lynch and Eilers (1991) found that 10–13-year-old children were better able to detect changes to the Western major than to the Javanese pelog scale, but this sensitivity likely emerges much earlier in development. Studies examining the ease of detection of within-key compared to out-of-key notes in Western melodies indicate that by 4 or 5 years of age, key membership for Western music is in place (Trehub et al. 1986; Trainor 2005; Corrigan and Trainor 2009).

Although musical scales do differ across musical systems, some structural properties are near universal, just as some properties of languages are near universal (Chomsky 1965). One common feature of musical scales is that they contain intervals of more than one size. These asymmetries of the interval structure are vital to musical pitch organization as they enable different notes of the scale to take on different functions. For example, in the Western major scale, every note forms a unique set of intervals or pitch relations with the other notes of the scale (Balzano 1980) so each takes on a different identity. Interestingly, even infants are better able to process unfamiliar scales with more than one interval size compared to unfamiliar scales with equal intervals (Trehub et al. 1999).

Asymmetrical interval structure enables scales to have *internal* structure as well, in the form of the tonal hierarchy. Different notes of the scale have different musical

meanings. The first note of the Western major scale, the tonic, is the most stable and compositions typically begin and end on the tonic. The second degree, or supertonic, sounds resolved if it falls to the tonic. Similarly, the last (seventh) scale degree, or leading tone, sounds resolved if it rises to the tonic. The fifth degree, or dominant, is the second most stable note of the scale and it bears a consonant relation (3:2 frequency ratio) to the tonic. The fourth degree, or subdominant, is the next most stable note, and also bears a consonant relation (4:3 frequency ratio) to the tonic. Together with the tonic and dominant, the third degree or mediant forms a tonic triad, a common melodic pattern. To examine the psychological reality of the tonal hierarchy Krumhansl developed the probe tone method, whereby a musical context such as a scale is presented, followed by a single probe tone, and participants rate how well the probe tone fits the context (Krumhansl 1990). Krumhansl found a consistent profile of ratings for the different notes of the Western scales (see Fig. 8.2) showing that Western adults are sensitive to musical key structure (Krumhansl and Cuddy 2010). Even adults without formal musical training are sensitive to these semantic relations (Trainor and Trehub 1994; Koelsch et al. 2002a, 2003a; Bigand and Poulin-Charronnat 2006). Using a similar probe-tone methodology, it has been shown that from first grade, children rate in-key notes as fitting better into the key context than out-of-key notes, but that sensitivity to the full tonal hierarchy does not emerge until later into the school-age years (Krumhansl and Keil 1982; Speer and Meeks 1985; Cuddy and Badertscher 1987).

Thus, young infants can remember melodies and are sensitive to universal aspects of scale structure such as unequal interval sizes. At the same time, it takes considerable experience for sensitivity to culture-specific scales to arise. Developmentally, sensitivity to key membership, or knowing which notes belong in a key and which do not, arises earlier than sensitivity to the tonal hierarchy, or understanding the different meanings of different notes within a key.

3.5 *Learning Harmonic Syntax*

Although particular musical scales differ across musical systems, the use of scales that divide the octave into a small number of discrete intervals is widespread. In contrast, the elaborate harmonic structure of Western music is fairly unique. In this regard, it is interesting that the implicit acquisition of harmonic rules is relatively late developing, likely not reaching adult levels until the early teenage years (Costa-Giomi 2003). Knowledge of harmonic structure depends on an underlying perception of consonance and dissonance. The structurally most important chords in a major key are based on the tonic (I chord), dominant (V chord), and subdominant (IV chord). (Roman numerals refer to the scale position of the note the chord is based on.) These are major chords, with the higher notes forming a major third (ratio of 5:4), perfect fifth (3:2) and octave (2:1) interval with the root or scale note of the chord. Chords built on the supertonic (ii chord), mediant (iii chord), and submediant (vi chord) are next most consonant, with the higher notes forming a minor third (6:5), perfect fifth (3:2) and octave (2:1) with the

root or scale note of the chord. The chord built on the leading tone is more dissonant, with the higher notes forming a minor third (6:5), diminished fifth (45:32) and octave (2:1) with the root of the chord. The sequential ordering of chords in composition follows probabilistic syntactic rules, such that certain chord progressions are much more likely than others. For example, V and vii° chords create tension, which is resolved when they are followed by a I chord. Tension and tone color can be enhanced with the addition of other notes to the chord, most commonly by the addition of a note that forms a seventh interval with the root of the chord. For example, when the minor seventh (16:9) interval is added to a dominant chord, it increases the dissonance and therefore the experienced tension, which in turn increases the sense of resolution when it is followed by a I chord. A diminished seventh is often added to a vii° chord, again to increase dissonance, tension, and degree of resolution.

Although an adult level of harmonic understanding is not achieved until much later, the beginnings of understanding syntax for chord progressions can be seen in preschool and school-age children. In an implicit task, Schellenberg et al. (2005) showed that children as young as 6 years of age were faster to identify features of the last chord of a sequence, such as its instrument timbre or sung phoneme, when it was an expected tonic (I) chord compared to when it was an unexpected subdominant (IV) chord. Using a child-friendly task involving puppets, Corrigan and Trainor (2009) showed that children as young as 4 years of age rated sequences as sounding “good” more often when they ended on an expected tonic chord compared to when they ended on an unexpected subdominant chord. Similar results have also been obtained with ERP measures. Children as young as 5 years of age showed different EEG responses when presented with a highly unexpected chord at the end a sequence of chords compared to an expected chord (Koelsch et al. 2003b).

Harmonic structure is so pervasive in Western music that even melodies presented in isolation set up expectations for an underlying harmonic chord-based accompaniment. The psychological reality of these implied harmonic expectations are apparent in adults’ detection of wrong notes in a melody. Trainor and Trehub (1994) showed that adults are better able to detect a wrong note in a melody that remains within the key of the melody when the wrong note violates the implied harmony expected at that point, compared to when it is consistent with that harmony. They found that 7-year-olds but not 5-year-olds were adult-like in this ability.

In sum, harmony is relatively rare across musical systems and sensitivity to harmonic structure develops relatively late. Although some sensitivity can be seen in preschool children, it continues to develop well into middle childhood.

3.6 Summary of Research on the Development of Musical Pitch Processing

Musical pitch structure is multifaceted. Certain basic aspects of pitch structure, such as pitch discrimination, consonance and dissonance, and transpositional invariance are processed very early in development. These abilities form the basis for the

development of higher-level pitch structures. Most important of these in Western music are scales (keys) and harmony (chord sequences). The use of scales as the basis of melodic composition is essentially universal across musical systems, and scales share certain properties such as octave equivalence, the use of a small number of notes or scale degrees per octave, and more than one size of interval between scale notes. These universal features of scales can all be processed in infancy. At the same time, the particular scales used vary considerably between musical systems and are learned through exposure some time after 8 months of age. Harmonic syntax, on the other hand, is relatively rare across musical systems and shows a protracted developmental trajectory. Sensitivity to some aspects of harmony appears by 4 years of age, but full sensitivity is probably not achieved until the teenage years.

4 Effects of Experience on Musical Pitch Processing

The general finding that development is greatly influenced by experience applies to musical acquisition (Trainor and Corrigan 2010). Just as young children learn to speak the language of their culture without formal training, they learn the pitch structure of their musical system as well, although adults without formal musical training typically do not have explicit knowledge of this structure. However, when implicit tasks are used, nonmusicians demonstrate quite sophisticated musical knowledge of key membership and harmonic syntax. In such tasks, participants are, for example, asked to make a judgment about whether the last chord of a sequence is in tune (Bharucha and Stoeckig 1986, 1987) or consonant (Bigand and Pineau 1997; Tillmann et al. 1998; Bigand et al. 1999). However, the manipulation of interest is whether or not the last chord is syntactically expected or not. Both musicians and nonmusicians are slower to make the above judgments when the chord is unexpected, indicating their knowledge of harmonic syntax. In the case of melodic processing, ERP studies indicate that even when melodic sequences are not attended to, the auditory cortex of nonmusicians responds to violations of expected pitch contour (up/down pattern of pitch changes) and expected melodic pitch interval (Trainor et al. 2002a).

Although nonmusicians develop networks in the brain that are specialized for processing the musical pitch structure they hear in their environment, there are vast differences between musicians and nonmusicians in the amount and type of musical experience they gain during development. Unlike language, where most people attend school and learn to read, typically musicians, but not nonmusicians, produce music, learn to read music and gain explicit knowledge about musical structure. In principle, these large differences in experience would be expected to affect brain development. Indeed, the brains of musicians and nonmusicians differ in many ways, as discussed later. However, in practice it is difficult to determine whether the differences seen in adulthood between musicians and nonmusicians are the result of differences in musical experience or whether they are largely due to genetic differences affecting early music processing and therefore the decision to train musically. The gold standard in this respect would be longitudinal studies with random assignment to groups receiving different musical experiences. In practice, this is difficult to

achieve, and there are few studies that meet this standard. Therefore, most of the evidence for effects of formal musical experience on development is suggestive rather than definitive.

4.1 Structural Brain Differences Between Musicians and Nonmusicians

Music recruits a wide range of cortical areas (e.g., Koelsch and Siebel 2005; Schlaug 2009). Two regions that appear to be particularly important for musical pitch are the superior temporal and inferior frontal areas. These areas and the white matter tracts connecting them are compromised in individuals with poor music pitch processing (Hyde et al. 2006, 2007; Loui et al. 2009). Interestingly, the network connecting these regions is critical for language as well as for music, and it is likely involved in the processing of patterns of sounds.

Differences between adult musicians and nonmusicians have been reported for a number of cortical regions using MRI measures. Musicians show increased gray matter in Broca's Area (Sluming et al. 2002), the cerebellum (Hutchinson et al. 2003), and motor areas (Gaser and Schlaug 2003; Bangert and Schlaug 2006). Further, the degree of volume increase in auditory cortex correlates with behavioral musical skill (Schneider et al. 2002). The thickness of cortex has also been reported to be greater in musicians compared to nonmusicians, particularly in the right secondary auditory cortex and dorsolateral prefrontal cortex (Bermudez et al. 2009).

Simply knowing that these differences between adult musicians and nonmusicians exist does not clarify the extent to which they are the result of genetic differences or of formal musical training and many hours of practice. There has only been one MRI study to compare young children taking music lessons with children not taking lessons (Hyde et al. 2009). Children were tested at 6 years of age and again 15 months later. There was no random assignment to groups, but half of the children took music lessons of some kind outside of school whereas the other half did not. Initially there were no differences between the groups. However, after the 15 months, the musician children showed larger volume in certain motor areas as well as in the right primary auditory cortex. Changes in the latter were correlated with changes in behavioral performance on melodic and rhythmic discrimination. In sum, this one study is consistent with musical experience causing structural changes in the brain that would enhance pitch processing. It remains unknown as to whether there is a critical period during which such experience would have the greatest effect (Trainor 2005).

4.2 Effects of Experience on Processing Single Tones

Unlike fMRI, which has temporal resolution on the order of seconds, electroencephalography (EEG) and magnetoencephalography (MEG) recordings can measure the stages of processing in the brain with a submillisecond resolution (Fig. 8.4), and

thus can provide detailed information about the processing of individual tones in the brain. Differences between adult musicians and nonmusicians have been found at virtually every stage of processing. At the subcortical level, musicians show larger and temporally less variable responses to sounds compared to nonmusicians (Musacchia et al. 2007; Kraus and Chandrasekaran 2010). They also show increased steady-state responses originating in primary auditory cortex (Schneider et al. 2002). Automatic preattentive responses N1b (e.g., Pantev et al. 1998), N1c, and P2 (e.g., Shahin et al. 2003), originating in secondary auditory cortices, are also larger in musicians. Later components are larger as well, including P3a responses reflecting inadvertent capture of attention and bringing of sounds into consciousness (Fujioka et al. 2004, 2005) and P3b, reflecting memory and conscious processing of sound (e.g., Besson and Faïta 1995; Trainor et al. 1999; Tervaniemi et al. 2005). It is difficult to prove that these differences are the result of experience, but several lines of research suggest that experience plays a large role. First, the size of the N1b (Pantev et al. 1998) and P3b (Trainor et al. 1999) responses are correlated with the age of onset of music lessons. Further, the increase in N1b appears to be specific to the timbre of the instrument of practice. Specifically, violinists show a larger increase in N1 for violin tones whereas trumpet players show a larger increase for trumpet tones (Pantev et al. 2001). Second, laboratory training studies show that the P2 response is particularly neuroplastic. Increases in P2 responses were seen after training on frequency discrimination only for tones at the frequency trained (Bosnyak et al. 2004). In addition, MMN amplitude increases after controlled laboratory training in learning to play melodies on a keyboard (Lappe et al. 2008).

There is little work on the effects of musical experience on the processing of isolated tones. However, Shahin and colleagues (2004) reported that N1 and P2 in response to single tones were larger in 4- to 5-year-old children engaged in Suzuki music lessons compared to children not training musically (Fig. 8.5b). Fujioka et al. (2006) measured MEG responses to musical tones and noises in similar-age children beginning music lessons, and in children engaging in other extracurricular activities, every 3 months for a year. The largest difference in how the two groups changed over the course of the year was in the N2 component to music sounds, but not to noises, which likely reflects increased automatic attention for musical tones with musical training.

Although the literature on the processing of isolated tones focuses on the various slow components described earlier, oscillatory activity reflecting the activation of networks of neurons can also be studied by analyzing EEG and MEG signals in the frequency domain. Induced activity in the gamma band range (30–100 Hz) is of interest as it is thought to reflect the operation of attention and memory and the interaction of top-down and bottom-up processes (e.g., Bertrand and Tallon-Baudry 2000). Induced gamma band activity occurs in response to a sound, but it is not precisely phase-locked to the sound stimulus. It is therefore thought to reflect the recruitment of ongoing oscillatory activity in the brain for the processing of the sound. Induced gamma band activity in response to musical tones has been shown to be larger in adult musicians than in nonmusicians (Shahin et al. 2008). Further, this activity was measured in response to violin, piano and pure tones in 4-year-old children at the onset of music lessons and 1 year later, and compared to a second

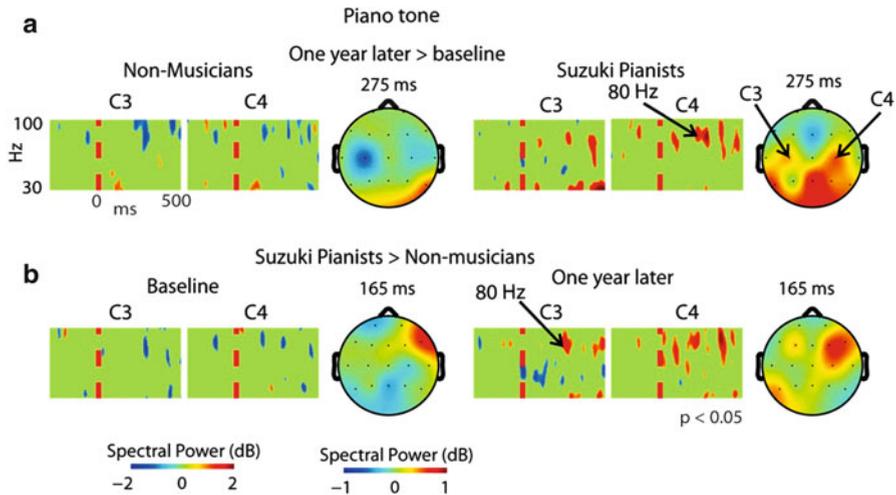


Fig. 8.8 Gamma band activity in 4- and 5-year-old children modulated by musical training. Induced gamma band activity (30–100 Hz) was derived from EEG recordings in response to piano tones measured in two groups of children (one beginning music lessons at the first measurement and the other not engaging in musical training) and at two time points 1 year apart (4 and 5 years of age). Zero ms represents the onset of piano tones. **(a)** Differences in gamma band activity across the year for the nonmusician children (*left*) and for the musician children (*right*). The musician children showed an increase in gamma band activity after a year, whereas the nonmusicians did not. As in adults, gamma band activity in response to a piano tone in musician children is long-lasting (at least 500 ms) and comes in waves, suggesting that it rides on a slower oscillatory process. **(b)** Differences in gamma band activity between the musician and nonmusician children at the first measurement (*left*) and at the second measurement a year later (*right*). At the first measurement there was no difference between groups, whereas by the second measurement, musician children showed greater gamma band activity compared to nonmusician children [From Trainor et al. (2009). Reprinted with permission of John Wiley & Sons]

group of children tested at the same two developmental time points (Fig. 8.8). None of the children showed significant gamma band activity at the first measurement. However, by the second measurement, gamma band activity was seen in the musician children but not in the control group (Shahin et al. 2008).

Together these studies of adult and child musicians suggest that formal musical training leads to the recruitment of more neurons and/or to increased synchronicity between neurons for the processing of isolated musical tones.

4.3 Effects of Experience on Melodic Processing

Melodies consist of sequences of tones and are composed within a musical system according to the tonal rules of that system. A sequence of melodic notes gives rise to expectations about what notes are more or less likely to follow. As discussed earlier, these expectations are learned through informal exposure to a musical system.

At the same time, there is substantial evidence that formal musical training greatly enhances melodic expectations.

More than half a decade ago, Meyer (1956) argued that the essence of musical meaning was in the degree to which musical expectations were realized or violated. In particular, meaningful violations set in motion physiological and psychological responses that give rise to emotional responses. One ERP component, the MMN, is generated in auditory cortex and reflects the operation of a mechanism for tracking expectations (see Picton et al. 2000; Näätänen et al. 2007; Trainor and Zatorre 2009). As described in Sect. 3.1, MMN occurs in response to virtually any type of sound change, including changes in pitch, duration, loudness, timbre, and spatial location. It also occurs, even in nonmusicians, in response to a change in a pattern of sounds, such as when occasional wrong notes are played in a melody that is presented in transposition to different starting notes (keys) from repetition to repetition (Trainor et al. 2002a).

MMN is typically of larger amplitude for melodic changes in adult musicians compared to nonmusicians (Fujioka et al. 2004). This enhancement for musicians extends to polyphonic contexts in which two melodies are played at the same time (Fujioka et al. 2005, 2008). Interestingly, for both musicians and nonmusicians, there was evidence that separate memory traces were formed in the brain for the two melodies, and that the representations were more robust for the higher-pitch melody, consistent with behavioral evidence of superior perception of the highest voice in polyphonic contexts (Crawley et al. 2002). While these studies suggest that formal musical training might enhance melodic processing in the brain, studies of young children are necessary in order to understand the relation between genetic and experiential factors.

There are few studies of the effects of musical training on melodic processing. However, Trainor et al. (2011) randomly assigned 4-month-old infants to one of two listening conditions. In one case, infants listened daily for a week to a CD of short children's melodies played in marimba timbre. In the other case, infants listened to the same melodies, but played in guitar timbre. Subsequently, ERPs were measured to occasional small changes in the pitch of a repeating tone when it was played in marimba timbre and when it was played in guitar timbre. They found that mismatch responses to the pitch change were larger for guitar than marimba tones only in the group exposed to guitar timbre at home, indicating that the musical experience led to more robust pitch responses for the familiarized timbre.

In sum, there is evidence of melodic processing differences between adult musicians and nonmusicians, but there are few studies in children to indicate the extent to which these differences are the result of musical training or due to genetic factors.

4.4 Effects of Experience on Acquisition of Key Membership and Harmony

As with cortical representations for isolated tones and for melodies, adult musicians show larger brain responses to syntactically unexpected chords (Koelsch et al. 2002b). Although the MMN response appears to be insensitive to key violations

(Trainor et al. 2002a; Fujioka et al. 2004, 2005), a similar response, the early right anterior negativity (ERAN) is sensitive to key and harmony violations in adults (Koelsch et al. 2001) even when factors such as amount of note repetition and degree of dissonance are controlled (Koelsch et al. 2007; Leino et al. 2007).

Three studies in children suggest that formal musical training accelerates acquisition of harmonic knowledge. Jentschke and Koelsch (2009) presented 11-year-old children with chord sequences and examined ERP responses when the last chord conformed to Western harmonic syntactic rules and when it did not. One group of participants were in the Saint Thomas Boys Choir in Leipzig and the other group were not training musically, although the two groups were matched for IQ and parents' education level. They found larger ERAN responses in the musician children, consistent with effects of musical training on the neural circuitry involved in harmonic processing, although of course the extent to which genetic differences contribute is not known.

Corrigall and Trainor (2009) examined two groups of 4- to 5-year-old children, one in which the children were just beginning music lessons and the other in which the children did not take lessons. Children were tested behaviorally at two time points, once before the onset of lessons and once approximately 8 months later. Children listened to sequences of five chords and judged whether each sequence sounded "good" or "bad." The sequences ended on either the tonic chord (i.e., conformed to Western harmonic structure), an out-of-key chord (tests key membership), or a chord that was within the key but not the expected harmony (tests harmonic syntax). Both groups showed some sensitivity to both key membership and harmonic syntax and the groups did not differ significantly at the first measurement. However, at the second measurement, the group taking music lessons showed superior performance, suggesting that early formal music lessons enhance harmonic processing.

Finally, there is evidence that musical training in infancy can accelerate knowledge of key membership (Gerry et al. 2010). In a comprehensive study, 6-month-old infants were randomly assigned to one of two types of music classes for a period of 6 months. The active music group received weekly lessons where parents and infants learned songs, moved to rhythms, and took turns with simple instruments such as little drums and xylophones. The passive group listened to Baby Einstein™ CDs in the background at their weekly lessons while infants and parents played at book, ball, block, and art stations. After 6 months of classes, when the infants were 12 months of age, they were given a battery of tests, including a test of sensitivity to musical keys. Infants were presented with two versions of an unfamiliar piano piece, one in its original form, and the other manipulated to alternate between two keys one semitone apart on every beat. This manipulation had the effect of making the second piece sound atonal because all 12 notes were present. Infants' preferences for the two versions were tested in a preference paradigm (see Sect. 3.2). The two groups differed significantly, with infants in the passive group showing no preference and infants in the active group preferring the original over the atonal version. Thus, when appropriate pedagogical approaches are used, music classes in infancy can accelerate acquisition of key membership. Because this study had random assignment to group, it can be inferred that the difference in musical experience caused the difference in the development of sensitivity to tonality. Thus, it can be

concluded that formal training does affect musical outcome and that this training can be effective early in development. However, it remains unknown as to whether there are critical periods during which certain types of musical experience have the greatest effects.

5 Summary

Music is ubiquitous across cultures, and serves important functions, including emotional communication between infants, aiding general cognitive development, and promoting social cohesion and cooperation between individuals who make music together. Musical pitch structure has many levels. Individual tones have an absolute pitch level and tones are perceived relative to each other along the dimensions of pitch height, tone chroma (octave equivalence), and consonance and dissonance. In this chapter, it has been argued that musical pitch structure builds on these domain-general aspects of auditory processing, and that these aspects operate in similar ways across the different musical systems in the world. From a developmental perspective, these aspects of pitch operate early in development and reflect a combination of innate structural constraints, such as the response of the inner ear, and learning constraints, such as the fact that tonotopic maps do not develop normally in the absence of structured sound input. These aspects of pitch processing are not specific to music, and operate similarly in speech.

There are also aspects of musical pitch structure that are unique to music. These include the use of discrete pitch and a small number of pitch intervals per octave that define sets of scale notes that are used for musical composition. The domain-specificity of scales can be seen in that pitch changes in speech tend to be continuous and the exact sizes of pitch excursions in linguistic utterances are not critical. At this level of musical pitch structure there is considerable variation across musical systems in terms of the particular scales (intervals) used, in the probabilistic rules by which notes are combined into melodies, and in the meanings of particular note combinations. Developmentally, musical scales must be learned, just as a specific language must be learned. This learning appears to occur fairly early. The research evidence suggests that children as young as 3 or 4 years certainly have implicit knowledge of key membership, and that with musical training, some understanding of tonality can be seen as early as 12 months of age.

Some musical systems also have harmonic pitch structure, in which particular combinations of notes are sounded simultaneously to produce chords that are built on the various notes of the scale. Chord sequences have a sort of statistical syntax, in that the transitional probabilities of one chord following another are quite stable within a musical system. This aspect of musical pitch structure does not appear to be universal and brain areas beyond auditory cortex are recruited for harmony processing. Developmentally, sensitivity to harmony appears fairly late. Preschool children show sensitivity to some aspects of harmony, but a full lay understanding (i.e., implicit knowledge in musically untrained adults) does not appear to be complete until the early teenage years.

There is evidence that all aspects of music pitch processing are affected by learning and experience but, given a normal auditory environment in which to develop, the more complex culturally specific aspects appear to be more dependent on experience. Future developmental research involving non-Western musical pitch structures and participants familiar with such structures is necessary to corroborate current evidence regarding cultural specificity and experience, which is largely limited to Western structures and listeners. At the same time, musically trained and musically untrained adults appear to process musical pitch similarly. The difference between the groups appears to be one of sensitivity. Adults with musical training show larger brain areas devoted to musical processing and larger and earlier event-related potentials in response to musical stimuli. Furthermore, these training effects can be observed in young children. Musical training studies with random assignment are rare, but provide the most direct evidence of the role of experience. Indeed there is evidence that the earlier musical training begins, the larger the plastic changes that are possible, and that such experientially driven change can be seen prior to 1 year of age. Given the role of music in social, emotional and cognitive processes, musical experience in the early years likely helps to set children on an optimal developmental trajectory.

Acknowledgments The writing of this chapter was supported by the Natural Sciences and Engineering Research Council of Canada, The Canadian Institutes of Health Research, the Canada Foundation for Innovation, and the Grammy Foundation. We thank Kathleen Corrigan for comments on an earlier draft.

References

- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing and early reading ability in preschool children. *Journal of Experimental Child Psychology*, 83, 111–130.
- Balzano, G. J. (1980). The group theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, 4, 66–84.
- Bangert, M., & Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *European Journal of Neuroscience*, 24, 1832–1834.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436, 1161–1165.
- Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, 19, 1583–1596.
- Bertrand, O., & Tallon-Baudry, C. (2000). Oscillatory gamma activity in humans: A possible role for object representation. *International Journal of Psychophysiology*, 38, 211–223.
- Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278–1296.
- Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 403–410.
- Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics. Special Issue: The Understanding of Melody and Rhythm*, 41, 519–524.

- Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics*, 59, 1098–1107.
- Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100, 100–130.
- Bigand, E., Madurell, F., Tillman, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 184–197.
- Bischoff Renninger, L., Wilson, M. P., & Donchin, E. (2006). The processing of musical scale: An ERP study of musicians trained outside of the western musical system. *Empirical Musicology Review*, 1, 185–197.
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated with reward and emotion. *Proceedings of the National Academy of Sciences of the USA*, 98, 11818–11823.
- Bosnyak, D. J., Eaton, R. A., & Roberts, L. E. (2004). Distributed auditory cortical representations are modified when nonmusicians are trained at pitch discrimination with 40 hz amplitude modulated tones. *Cerebral Cortex*, 14, 1088–1099.
- Burns, E. M. (1999). Intervals, scales and tuning. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 215–264). San Diego: Academic Press.
- Chang, H. W., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, 24, 324–331.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clarkson, M. G., & Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of the Acoustical Society of America*, 77, 1521–1528.
- Clarkson, M. G., & Rogers, E. C. (1995). Infants require low-frequency energy to hear the pitch of the missing fundamental. *Journal of the Acoustical Society of America*, 98, 148–154.
- Corrigall, K., & Trainor, L. J. (2009). Effects of musical training on key and harmony perception. *Annals of the New York Academy of Sciences*, 1169, 164–168.
- Costa-Giomi, E. (2003). Young children’s harmonic perception. *Annals of the New York Academy of Sciences*, 999, 477–484.
- Crawley, E. J., Acker-Mills, B. E., Pastore, R. E., & Weil, S. (2002). Change detection in multi-voice music: The role of musical structure, musical training, and task demands. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 367–378.
- Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics*, 41, 609–620.
- Demany, L., & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America*, 76, 57–66.
- Deutsch, D., Henthorn, T., & Dolson, M. (2004). Absolute pitch, speech and tone language: Some experiments and a proposed framework. *Music Perception*, 21, 339–356.
- Fishman, Y. I., Reser, D. H., Arezzo, J. C., & Steinschneider, M. (1998). Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey. *Brain Research*, 786, 18–30.
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, 16, 1010–1021.
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and non-musicians. *Journal of Cognitive Neuroscience*, 17, 1578–1592.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain: A Journal of Neurology*, 129, 2593–2593.
- Fujioka, T., Trainor, L., & Ross, B. (2008). Simultaneous pitches are encoded separately in auditory cortex: An MMN study. *NeuroReport*, 19, 361–366.
- Fujioka, T., Mourad, N., & Trainor, L. J. (2011). Development of oscillatory neural networks and the auditory-specific tau rhythm in infancy. *European Journal of Neuroscience*, 33, 521–529.

- Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, 23, 9240–9245.
- Gerry, D., Unrau, A. J., & Trainor, L. J. (2010). Participation in active infant music classes accelerates acquisition of scale structure knowledge. Presented at the International Conference on Music Perception and Cognition, August 23–27, 2010, Seattle, WA.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19, 893–906.
- He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *The Journal of Neuroscience*, 29, 7718–7722.
- He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, 19, 878–892.
- He, C., Hotson, L., & Trainor, L. J. (2009). Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. *Neuropsychologia*, 47, 218–229.
- Huron, D. (2003). Is music an evolutionary adaptation? In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 57–75). Oxford: Oxford University Press.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: The MIT Press.
- Hutchinson, S., Lee, L. H., Gaab, N., & Schlaug, G. (2003). Cerebellar volume of musicians. *Cerebral Cortex*, 13, 943–949.
- Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain*, 129, 2562–2570.
- Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: When less is better than more. *The Journal of Neuroscience*, 27, 13028–13032.
- Hyde, K. L., Lerch, J. P., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience*, 29, 3019–3025.
- Ilari, B., & Polka, L. (2006). Music cognition in early infancy: Infants' preferences and long-term memory for Ravel. *International Journal of Music Education*, 24, 7–20.
- Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, 19, 2579–2594.
- Janata, P., Birk, J. L., Van Horn, J. D., Lemmon, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, 298, 2167–2170.
- Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, 47, 735–744.
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, 31, 354–364.
- Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9, S78–S84.
- Koelsch, S., Gunter, T. C., & Friederici, A. D. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 13, 520–541.
- Koelsch, S., Gunter, T. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*, 12, 1385–1389.
- Koelsch, G., Gunter, T. C., Yves, D., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002a). Bach speaks: A cortical “language-network” serves the processing of music. *NeuroImage*, 17, 956–966.
- Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002b). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, 39, 657–663.
- Koelsch, S., Gunter, T., Schröger, E., & Friederici, A. D. (2003a). Processing tonal modulations: An ERP study. *Journal of Cognitive Neuroscience*, 13, 520–541.
- Koelsch, S., Grossman, T., Gunter, T. C., Hahne, A., Schröger, E., & Friederici, A. D. (2003b). Children processing music: Electric brain responses reveal musical competence and gender differences. *Journal of Cognitive Neuroscience*, 15, 683–693.

- Koelsch, S., Jentschke, S., Sammler, D., & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, 44, 476–490.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11, 599–605.
- Krumhansl, C. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.
- Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception*. New York: Springer.
- Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition*, 10, 243–251.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89, 334–368.
- Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience*, 28, 9632–9639.
- Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, 1142, 169–177.
- Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *Journal of Neuroscience*, 29, 10215–10220.
- Lynch, M. P., & Eilers, R. E. (1991). Children's perception of native and nonnative musical scales. *Music Perception*, 9, 121–132.
- Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, 1, 272–276.
- Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18, 199–211.
- Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science*, 9, 46–50.
- McDermott, J., & Hauser, M. D. (2005). The origins of music: Innateness, uniqueness and evolution. *Music Perception*, 23, 29–59.
- McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, 20, 1035–1041.
- Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press: Chicago.
- Moreno, S., & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology*, 43, 287–291.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the USA*, 104, 15894–15898.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118, 2544–2590.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392, 811–814.
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport*, 12, 169–174.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36, 767–776.
- Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, 24, 6810–6815.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308, 1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533–546.
- Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology & Neurotology. Special Issue: Mismatch Negativity*, 5, 111–139.

- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, 98, 1–11.
- Plantinga, J., & Trainor, L. J. (2008). Infants' memory for isolated tones and the effects of interference. *Music Perception*, 26, 121–127.
- Plantinga, J., & Trainor, L. J. (2009). Melody recognition by two-month-old infants. *Journal of the Acoustical Society of America*, 125, EL58–62.
- Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 38, 548–560.
- Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, 111, 220–236.
- Ross, D. A., Olson, I. R., Marks, L. E., & Gore, J. C. (2004). A nonmusical paradigm for identifying absolute pitch possessors. *The Journal of the Acoustical Society of America*, 116, 1793–1799.
- Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, 37, 74–85.
- Saffran, J. R., Loman, M. M., & Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, 77, 15–23.
- Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science*, 14, 322–325.
- Schellenberg, E. G., & Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *Journal of the Acoustical Society of America*, 100, 3321–3328.
- Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, 14, 262–266.
- Schellenberg, E. G., Bigand, E., Poulin-Charronnat, B., Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in western music. *Developmental Science*, 8, 551–566.
- Schlaug, G. (2009). Music, musicians and brain plasticity. In S. Hallen, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 197–207). Oxford: Oxford University Press.
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, 5, 688–694.
- Schönwiesner, M., & Zatorre, R. J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, 187, 97–105.
- Shahin, A., Bosnyak, D., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, 23, 5545–5552.
- Shahin, A., Roberts, L. E., & Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *NeuroReport*, 15, 1917–1921.
- Shahin, A. J., Roberts, L. R., Chau, W., Trainor, L. J., & Miller, L. (2008). Musical training leads to the development of timbre-specific gamma band activity. *NeuroImage*, 41, 113–122.
- Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2346–2353.
- Shepard, R. N. (1965). Approximation to uniform gradients of generalization by monotone transformations of Scale. In D. I. Mostofsky (Ed.), *Stimulus generalization*. Stanford, CA: Stanford University Press
- Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89, 305–333.
- Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, 17, 1613–1622.
- Smith, N. A., & Cuddy, L. L. (2003). Perceptions of musical dimensions in Beethoven's Waldstein Sonata: An application of tonal pitch space theory. *Musicae Scientiae*, 7, 7–34.
- Speer, J. R., & Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology*, 5, 49–56.

- Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, 55, 1061–1069.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch-discrimination accuracy in musicians vs. non-musicians—an event-related potential and behavioral study. *Experimental Brain Research*, 161, 1–10.
- Tew, S., Fujioka, T., He, C., & Trainor, L. (2009). Neural representation of transposed melody in infants at 6 months of age. *Annals of the New York Academy of Sciences*, 1169, 287–290.
- Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 156, 203–227.
- Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, 61, 157–174.
- Tillmann, B., Bigand, E., Escoffier, N., & Lalitte, P. (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology*, 18, 343–358.
- Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed play songs and lullabies. *Infant Behavior and Development*, 19, 83–92.
- Trainor, L. J. (1997). The effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1427–1438.
- Trainor, L. J. (2005). Are there critical periods for music development? *Developmental Psychobiology*, 46, 262–278.
- Trainor, L. J. (2008a). The neural roots of music. *Nature*, 453, 598–599.
- Trainor, L. J. (2008b). Event-related potential (ERP) measures in auditory developmental research. In L. A. Schmidt & S. J. Segalowitz (Eds.), *Developmental psychophysiology: Theory, systems and methods* (pp. 69–102). New York: Cambridge University Press.
- Trainor, L. J., & Corrigan, K. A. (2010). Music acquisition and the effects of musical experience. In M. R. Jones, R. R. Fay & A. N. Popper (Eds.), *Music perception*. New York: Springer.
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior & Development*, 21, 77–88.
- Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 394–402.
- Trainor, L. J., & Trehub, S. E. (1993). What mediates adults' and infants' superior processing of the major triad? *Music Perception*, 11, 185–196.
- Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics*, 56, 125–132.
- Trainor, L. J., & Zacharias, C. A. (1998). Infants prefer higher-pitched singing. *Infant Behavior and Development*, 21, 799–805.
- Trainor, L. J., & Zatorre, R. (2009). The neurobiological basis of musical expectations: From probabilities to emotional meaning. In S. Hallen, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 171–182). Oxford: Oxford University Press.
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, 20, 383–396.
- Trainor, L. J., Desjardins, R. N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology Special Issue: Music as a Brain and Behavioural System*, 51, 147–153.
- Trainor, L. J., McDonald, K. L., & Alain, C. (2002a). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, 14, 430–442.
- Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002b). Preference for consonance in 2- and 4-month-old infants. *Music Perception*, 20, 187–194.
- Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, 7, 289–296.
- Trainor, L. J., Shahin, A., & Roberts, L. E. (2009). Understanding the benefits of musical training: Effects on oscillatory brain activity. *Annals of the New York Academy of Sciences*, 1169, 133–142.

- Trainor, L. J., Lee, K., & Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: Specific effects of experience with musical timbres. *Brain Topogr*, 24, 192–203.
- Tramo, M. J., Cariani, P. A., Delgutte, B., & Braidia, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. In R. J. Zatorre, & I. Peretz (Eds.), *The biological foundations of music* (pp. 92–116). New York: New York Academy of Sciences.
- Trehub, S. E. (2009). Music lessons from infants. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 229–234). Oxford: Oxford University Press.
- Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and playsongs. *Advances in Infancy Research*, 12, 43–77.
- Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821–830.
- Trehub, S. E., Cohen, A. J., Thorpe, L. A., & Morrongiello, B. A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 295–301.
- Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 965–975.
- Vaughn, K. (2000). Music and mathematics: Modest support for the oft-claimed relationship. *Journal of Aesthetic Education*, 34, 149–166.
- Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science*, 9, 583–589.
- Wallin, N. L., Merker, B., & Brown, S. (Eds.) (2000). *The origins of music*. Cambridge, MA: MIT Press.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8, 547–558.
- Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior & Development*, 21, 483–492.

Chapter 9

Atypical Auditory Development and Effects of Experience

Laurie S. Eisenberg, Karen C. Johnson, Sophie E. Ambrose,
and Amy S. Martinez

1 Introduction

Early onset of permanent hearing loss results in atypical auditory development, a condition with known deleterious effects for spoken language acquisition, psychosocial development, and academic accomplishment. The extent of these consequences varies according to a number of factors, of which magnitude of hearing loss and age at time of identification/intervention are two of the most prominent. Variability in outcomes also is attributable to the effectiveness of an auditory sensory device (hearing aid or cochlear implant or both) and quality of the intervention program, including communication mode.

There is converging evidence both from the psychoacoustic and speech perception literature to suggest that auditory perceptual capabilities continue to develop as the child with normal hearing matures (Buss, Hall, and Gross, Chap. 4; Leibold, Chap. 5; Litovsky, Chap. 6; Panneton and Newman, Chap. 7), reinforced by the anatomy and physiology of the developing auditory system (Keefe and Abdala, Chap. 2; Moore and Eggermont, Chap. 3). This chapter highlights the prominent research findings related to pediatric hearing loss from a developmental perspective. Attention is given to degree of hearing loss, age at identification, and communication mode/intervention with emphasis on speech perception and production. Research findings relevant to the effects of auditory deprivation and experience follow.

By way of an introduction, some basic information about the way in which hearing loss is characterized will be useful in understanding concepts presented throughout the chapter. There are three broad categories of hearing loss: conductive, sensorineural, and mixed (a combination of conductive and sensorineural). Conductive hearing

L.S. Eisenberg (✉) • K.C. Johnson • S.E. Ambrose • A.S. Martinez
Children's Auditory Research & Evaluation Center, House Research Institute,
Los Angeles, CA, USA
e-mail: leisenberg@hei.org; kjohnson@hei.org; sophie.ambrose@boystown.org; amartinez@hei.org

loss relates to pathology of the outer and middle ear in which the acoustic signal is conveyed or “conducted” to the sensory mechanism (cochlea or inner ear). Middle-ear pathologies typically can be remediated using medical or surgical intervention. Sensorineural hearing loss relates to pathology of the cochlea or eighth nerve. Specialized objective measures, such as otoacoustic emissions and auditory brain stem responses, can differentiate between cochlear and neural pathology. Sensory and neural deficits cannot be restored to normal functioning using medical or surgical treatment, thus sensory assistance is recommended. The research summarized in this chapter is specific to permanent, sensorineural hearing loss.

Degree and configuration are useful descriptors for characterizing hearing loss. Degree of hearing loss is routinely classified by severity or magnitude of loss (ranging from mild to profound), whereas configuration of loss depicts the shape of the audiometric pattern across frequencies (flat, gently sloping, steeply sloping, reverse slope, etc.). Degree of hearing loss is often represented by the average across three test frequencies (500, 1,000, and 2,000 Hertz, or Hz), also termed the pure-tone average (PTA). There are slight variations in the way degree of hearing loss may be delineated. For the purpose of this chapter, the following designations are used: slight, 16–24 dB hearing level (dB HL); mild, 25–39 dB HL; moderate, 40–54 dB HL; moderately severe, 55–69 dB HL; severe, 70–89 dB HL; profound, 90 dB HL and greater. Whereas hearing loss configuration often is not flat across frequencies, descriptors may span two or more hearing loss classifications, such as in mild to moderate.

2 Pediatric Hearing Loss and Sensory Assistance

The most immediate and direct consequence of permanent hearing loss is the diminished capacity to perceive and interpret sound. These skills range from basic sound awareness to higher level processing, which enable the listener to detect, discriminate, identify, and comprehend auditory stimuli. Hearing is directly linked to the development of spoken language. For adults or older children, who have had the opportunity to develop spoken language before the onset of hearing loss, receptive communication is hindered by the loss. In contrast, prelingual onset of hearing loss is much more debilitating than postlingual loss because it compromises the individual’s ability to acquire spoken language, both receptively and expressively. Infants must be able to attend selectively to speech to extract information about the organization of sound pattern information in the speech signal. It has been suggested that the absence of auditory stimulation during the period of early neural development *in utero* and very early childhood may impair the child’s ability to attend to the properties of speech that lay the foundation for speech discrimination and word learning skills (Houston 2002; Houston et al. 2003a). It is for this reason that the appropriate selection of an auditory sensory device as early as possible is essential to the development of auditory skills in the child with hearing loss.

The primary goal of sensory assistance is to ensure audibility across a broad range of frequencies. Although the unassisted or “unaided” listening condition in an ear is used to classify the magnitude of the hearing loss, evidence of suprathreshold

performance is often measured in the sound field with the child's sensory device activated. The two primary types of sensory devices in use today are hearing aids and cochlear implants. Hearing aids selectively amplify the acoustic signal as a function of frequency to accommodate both severity and configuration of the hearing loss. In contrast, cochlear implants bypass the damaged cochlear structures and directly stimulate residual neural elements via electrical stimulation. Each of these devices relies on complex digital signal processing for individualized custom fitting.

Children with unaided mild to severe hearing loss generally benefit from hearing aids, which are likely to amplify much of the speech spectrum into the audible range. Even so, depending on severity and configuration of the hearing loss or the gain and bandwidth of the hearing aid, the high-frequency spectral cues of speech may continue to be inaudible (Boothroyd and Medwetsky 1992; Stelmachowicz et al. 2001, 2002). The inability to perceive high-frequency phonemes, such as /s/ or /z/, has implications for acquiring morphological structures in language, such as the plural form (Elfenbein et al. 1994). For milder degrees of loss, the speech signal may be of sufficient intensity without sensory assistance when in close proximity, such as the situation of a mother holding her infant. However, even children with mild hearing loss require sensory assistance for consistent audibility under diverse listening situations.

With hearing losses that are equal to or exceed 90 dB HL, it is often not possible to provide sufficient acoustic amplification to access much of the spectral information that is necessary for adequate speech recognition. Moreover, as hearing loss increases beyond 110 dB HL, responses from the hearing aid are more likely to be based upon vibrotactile stimulation than actual audition (Boothroyd and Cawkwell 1970; Erber 1972). For children with profound hearing loss, a cochlear implant typically is recommended. The cochlear implant enables the child with profound hearing loss, on average, to function similarly to a child with a severe loss (Blamey et al. 2001; Eisenberg et al. 2004). In recent years, there has been a trend toward implanting children with a greater amount of residual hearing than was once previously considered. There is evidence to indicate that preimplant average hearing thresholds better than 95 dB HL result in higher postimplant speech perception scores, at least within the first year of cochlear implant use (Gordon et al. 2001). However, because of the broad variability within hearing loss categories, there may be some individual children with severe or profound hearing loss who perform better with hearing aids than what might be expected from use of an implant. Frequency compression in hearing aids is now providing better access to high frequencies compared to the more traditional devices (Scollie et al. 2007). Although good clinical practice dictates that each child be evaluated with hearing aids before consideration of a cochlear implant, recent evidence also shows that a lengthy trial with hearing aids for implant candidates may have negative consequences for acquisition of spoken language (Niparko et al. 2010).

Complicating the picture is asymmetrical hearing loss, in which one ear has a profound hearing loss while the other ear has aidable residual hearing. Today, many centers will implant children with significant residual hearing in one ear, thereby pairing a hearing aid in the better hearing ear and a cochlear implant in the poorer hearing ear (Holt et al. 2005). This configuration is, at times, advisable because use of a hearing aid

in the ear opposite to the implant may provide supplemental low-frequency acoustic cues, even for children with bilateral severe to profound hearing loss (Straatman et al. 2010). The combination of an implant in one ear and a hearing aid in the other ear is referred to as a bimodal fitting. There also is a growing trend for the implantation of both ears, referred to as bilateral cochlear implantation. Children are either implanted sequentially (two different surgeries) or simultaneously (two devices implanted during the same surgery). For both bimodal and bilateral device configurations, the goal is the same as for fitting two hearing aids—to provide binaural hearing, which includes the ability to localize sound and to improve speech recognition in noise. Results to date suggest that some children are able to localize sound with bimodal or bilateral listening (Ching et al. 2007). In terms of speech recognition in noise, many children have shown improved performance in the bimodal or bilateral listening condition relative to the monaural condition, although in some cases there is no difference in performance (Ching et al. 2001; Holt et al. 2005; Litovsky et al. 2006). Much more detailed information about binaural hearing can be found in Litovsky, Chap. 6.

With the appropriate fitting of a sensory device, the child with hearing loss has access to an audible signal. The extent of this access is determined in large part by degree of hearing loss, as discussed in the next section.

3 Degree of Hearing Loss

Conventional wisdom would hold that auditory capabilities decrease as the level of hearing loss increases in magnitude. Although this finding is generally true, cochlear implantation of children with profound hearing loss is changing this view. The importance of degree of hearing loss on the development of speech perception and production in children is summarized here.

3.1 *Speech Perception*

As infants mature, they learn to associate the sounds they hear with the objects and events that generate those sounds, including the sounds of speech (Boothroyd 1997). Speech percepts become internally represented as linguistic units (i.e., syllables, words, phrases, sentences). The ability to recognize such units relies in part on the individual's capacity to code the intensity, temporal, and frequency (spectral) cues of speech (Boothroyd 1978). Much of the acoustic-phonetic information required for differentiating vowels from consonants or for processing the features of consonant voicing and manner are dependent on time and intensity cues, which are generally accessible over a wide range of frequencies. The prosodic or suprasegmental speech features (intonation, stress, duration, and number of syllables) also rely on time and intensity information, although there is some dependence on frequency for encoding intonation. In contrast, place of articulation is primarily reliant on frequency cues, and is restricted to the higher frequencies.

Sensorineural hearing loss results in reduced audibility, for which the high frequencies are typically the most susceptible. With increasing severity of hearing loss, cochlear processing of spectral, temporal, and intensity information becomes impaired (Moore 1996). The loss of audibility in combination with cochlear distortions would predict a diminution of speech perception abilities. Such an association was shown by Boothroyd (1984) in a study performed on 120 orally trained adolescents with PTA losses ranging between 55 and 123 dB HL. The children were assessed on speech contrast perception and open-set phoneme recognition. Strong relationships were shown between PTA and phoneme recognition ($r = -0.78$), and between PTA and speech contrast composite score ($r = -0.78$). In terms of speech contrast perception, the children with hearing loss ranging between 55 and 70 dB HL had access to all the speech contrasts assessed. Consonant place was the contrast most susceptible to hearing loss—scores decreased to 50% for losses of 75 dB HL and greater. The order by which scores fell to 50% for the remaining contrasts were shown for increasing levels of hearing loss: initial consonant continuance (85 dB HL), initial consonant voicing (90 dB HL), vowel place (100 dB HL), talker sex (105 dB HL), syllabic pattern (115 dB HL), and vowel height (beyond 115 dB HL). Particularly noteworthy was the finding that children with the most profound hearing loss were able to access several cues of speech, primarily those cues that relied on time and intensity information. Boothroyd surmised that the most intensive speech peaks in the speech signal were being accessed by these listeners.

Much of the original research published on the speech perception abilities of children with hearing loss has been of clinical origin and predominantly focused on older children. There are several possible explanations for the paucity of data on very young children with hearing loss. One reason is that the early studies were published before the widespread practice of newborn hearing screening, in which children were typically identified with hearing loss at the toddler stage (or even older) with additional lags for sensory device fitting (Mace et al. 1991; Harrison and Roush 1996). Also, many of the speech perception tests available for this population are composed of word stimuli, requiring that children have the requisite vocabulary to be administered such tests. The speech perception results from these clinical studies either describe the population on the basis of age or degree of hearing loss or both (e.g., Davis et al. 1986; Worthington et al. 1986; Wake et al. 2004), or explain the impact of such loss on language, speech production, or other developmental domains (Osberger 1986; Levitt et al. 1987; Elfenbein et al. 1994). Comparing results across studies becomes complicated because of variations in speech perception tests and protocols (open set or closed set; live voice or recorded), stimuli (phonemes, syllables, words or sentences), transducers (earphones or loudspeaker), and listening conditions (quiet or various background competitors). Despite these variations across protocols, general trends from several representative studies can be cited, including studies that utilized either closed-set word identification or open-set word recognition tests.

Noteworthy is the investigation by Davis et al. (1986), in which identification skills were evaluated in 40 school-age children (ages 5–18 years) with mild to severe hearing loss. Testing was administered under headphones (unaided scores) and in the sound field with hearing aids activated. A closed-set word identification test, the Northwestern

University—Children’s Perception of Speech (NU-CHIPS, Elliott and Katz 1980), was administered in quiet and in noise. Scores in the unaided listening condition ranged from 56% to 94%. With hearing aids activated, scores improved to 91% and higher. The amplified scores remained relatively high (72–87%) when the test was administered in low-level noise presented at 10 dB signal-to-noise ratio (S/N).

In a study focused on children with either mild to moderate or steeply sloping hearing loss, Hicks and Tharpe (2002) assessed the open-set word recognition abilities of 14 school-age children (6–11 years of age) with losses ranging from 25 to 70 dB HL. The test administered was the widely used Phonetically Balanced Word Lists—Kindergarten (PBK, Haskins 1949), presented in quiet and in background babble. With hearing aids activated, performance scores averaged 85% or better both in the quiet condition and in 20-talker speech babble (+20, +15, and +10 dB S/N). These results demonstrate that children with fairly modest hearing loss, particularly in the low and mid frequencies, perform quite well with acoustic amplification even in the presence of low-level background noise.

PBK word recognition was investigated by Eisenberg et al. (2004) in two samples of children with hearing loss; 25 school-age children (mean age=8.9 years) with moderately severe to profound hearing loss who used hearing aids, and 45 school-age children (mean age=7.7 years) with profound hearing loss who used cochlear implants. The children were further stratified on the basis of communication mode (oral and total communication). With hearing aids activated, the mean word score for 20 orally communicating children was 52%. In comparison, the mean word score for the five children who used total communication (oral and sign language combined) was 22%. The 33 orally communicating children with cochlear implants achieved a mean score of 60% with their devices activated, which was slightly higher than the mean score of the children with hearing aids. The 12 children with implants who used total communication obtained a higher score (mean=41%) than their hearing-aid comparison peers, although they remained below that of the orally communicating implant group.

The recent innovations in state-mandated newborn hearing screening have necessitated development of new assessment tools for infants and toddlers with hearing loss that can be applied in the clinical setting. The speech perception abilities of infants with normal hearing have been documented in a number of investigations using stringent assessment protocols (see Panneton and Newman, Chap. 7). Similar procedures are now being applied to infants and toddlers with hearing loss, and in most cases for those with cochlear implants (Houston et al. 2003a; Barker and Tomblin 2004; Uhler et al. 2011).

The challenges involved in assessing speech perception in infants with normal hearing are compounded when testing infants and toddlers with varying degrees of hearing loss. The approach taken by investigators at the House Research Institute has been to apply a conditioned head-turn response in assessing the infant’s ability to detect differences between vowel-consonant-vowel (VCV) stimuli (e.g., oodoo vs. eedee for vowel place). The test is called the VRASPAC (Visual Reinforcement of the Perception of Speech Pattern Contrasts; Eisenberg et al. 2007; Martinez et al. 2008). Derived from probability theory, the analysis algorithm computes the percent confidence that the child’s head-turn response is not random. Figure 9.1 displays test profiles of four

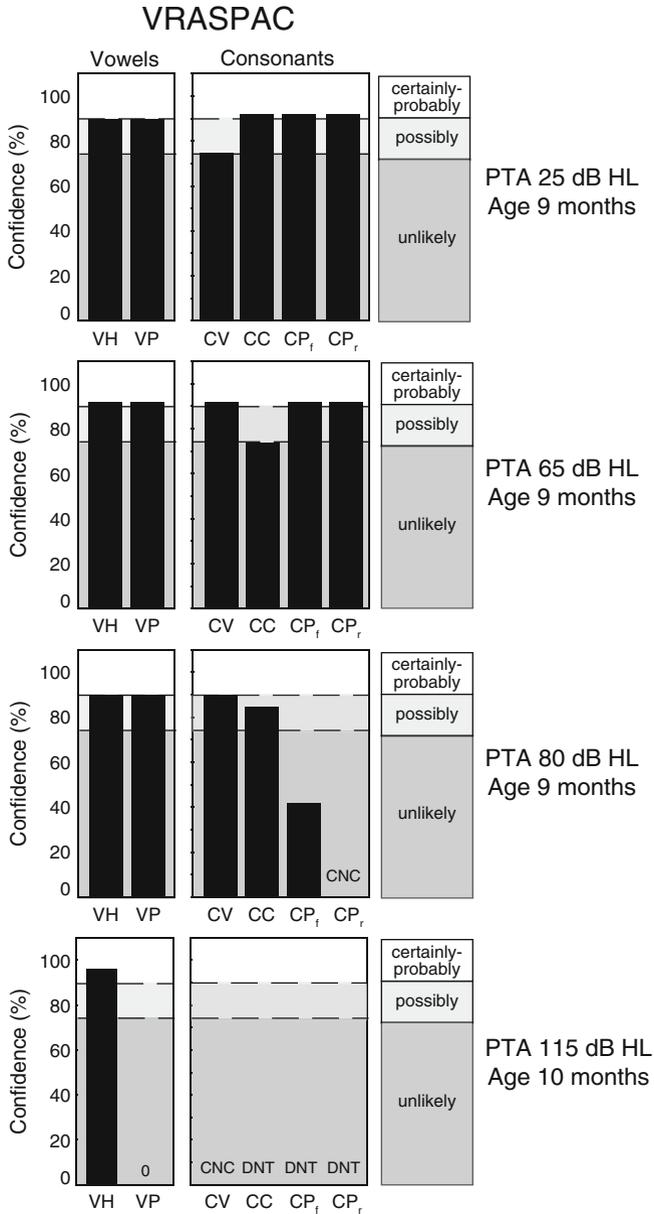


Fig. 9.1 VRASPAC results on four infants with different degrees of hearing loss assessed with hearing aids activated. The bars represent percent confidence for the following speech contrasts assessed: vowel height (VH), vowel place (VP), consonant voicing (CV), consonant continuance (CC), front consonant place (CP_f), and rear consonant place (CP_r)

9-month-old infants with different degrees of hearing loss, each of whom was assessed in the sound field with hearing aids activated. The speech features assessed were vowel height (VH) and place (VP), and consonant voicing (CV), continuance (CC), front place (CPf), and rear place (CPr). Systematic decrements can be seen in performance as a function of increasing hearing loss. The two children with the better hearing (PTA 25 and 65 dB HL) demonstrated high percent confidence across most of the contrasts assessed. The child with the least amount of residual hearing (PTA 115 dB HL) was successfully assessed only on the vowel height contrast. Although these results represent single cases, these data are relatively consistent with the mean data shown for the large sample of older children with hearing loss in Boothroyd (1984).

The relationship between degree of hearing loss and speech perception abilities has become blurred since the availability of cochlear implants. This finding was recently reinforced by Sininger et al. (2010). They reported that the association between unaided degree of hearing loss and speech perception outcomes was *not* statistically significant in a sample of children with unaided hearing loss ranging from mild to profound who were assessed with their sensory devices activated (hearing aids or cochlear implants).

Another factor that may counteract the effects of severity of loss is multimodal processing of speech. Children with hearing loss rely on visual cues to supplement speech sounds that may not be perceptible by audition alone. The addition of the visual modality can resolve perceptual ambiguities in the acoustic signal for the child with hearing loss, particularly when listening under adverse listening conditions such as noise or competing speech. The effects of auditory-visual processing have been described for children with hearing loss (Hack and Erber 1982; Lachs et al. 2001; Bergeson et al. 2005). Results across studies generally demonstrate that speech recognition is more accurate when hearing-impaired children are assessed with hearing plus lipreading compared to hearing only or lipreading only.

3.2 *Speech Production*

Early onset hearing loss adversely affects the child's ability to articulate the different sounds in his or her phonology. Although speech production skills are known to improve as the child matures, the extent of improvement will be counteracted by increasing magnitude of hearing loss. Severity of hearing loss plays an important role in early speech sound acquisition; children with milder degrees of hearing loss demonstrate larger phonetic and syllabic repertoires than do children with more severe losses (Yoshinaga-Itano and Sedey 2000; Davis et al. 2005; Moeller et al. 2007a). Consequently, acquisition of first words is likely to be delayed even among infants who were early identified and fitted with sensory devices within the first year of life (Moeller et al. 2007b; Nott et al. 2009).

Degree of hearing loss has been found to be a significant predictor of speech sound acquisition in the developing child. In a longitudinal study of vocal behavior in young children with hearing loss, Moeller et al. (2007a) observed that, as a group,

even infants identified on average at 2.5 months and amplified on average at 4.5 months tended to lag behind typically hearing age-mates in phonetic development over the first 24 months of life. Nonetheless, infants with PTA less than 50 dB HL tended to achieve consistent canonical babbling earlier than infants with more severe degrees of hearing loss. Canonical babbling, or the combining of consonant vowel or vowel consonant syllables in repetitive sequences, is developmentally significant because it lays the foundation for producing the first spoken words (Oller et al. 1975; Vihman et al. 1986). It is noteworthy that cochlear implantation has had a positive influence on the vocal development of young children with profound hearing loss in that babbling behavior has been documented at 1–6 months following device activation (Ertmer and Mellon 2001; Schauwers et al. 2004; Moeller et al. 2007a).

Yoshinaga-Itano and Sedey (2000) found significant differences in the phonetic repertoires and ratings of speech intelligibility for young children between 14 and 60 months of age grouped according to degree of loss. Although systematic differences in median number of vowel and consonant types were observed as a function of hearing loss group through approximately 42 months of age, by 60 months of age the children with PTA between 56 and 90 dB HL had caught up with those with PTA between 26 and 55 dB HL in number of phonetic elements produced. The more severely impaired infants were rated as nearly as intelligible as those with milder loss. No significant relationship was found between age of identification (<6 vs. >6 months) and speech production outcome. Cochlear implantation has enabled the expansion of phonetic inventories, with some children attaining vocal milestones with fewer months of auditory experience than children with normal hearing (Ertmer and Mellon 2001; Ertmer et al. 2007).

Taken together, these results provide evidence to reinforce the notion that increasing severity of hearing loss for the most part results in decreasing speech perception/production skills in children during a critical time when they are learning spoken language. Use of a sensory device enables many inaudible acoustic cues to be detected. With the use of hearing aids, a child with a moderately severe loss may function more like a child with a mild hearing loss. With use of a cochlear implant, a child with a profound hearing loss may function like a child with a severe hearing loss. Because of recent developments in newborn hearing screening coupled with earlier fitting of an auditory sensory device, opportunities have arisen to investigate whether auditory developmental trajectories are influenced when the duration of auditory deprivation is reduced. The following section addresses the effects of age at identification.

4 Age at Identification of Hearing Loss

Of the many factors impacting development in children with hearing loss, age at identification has long been regarded as among the most important (Davis 1965; Downs 1970). Early attempts to implement newborn hearing screening targeted

those populations considered to be at high risk for hearing loss for a variety of reasons, including family history, congenital malformations or stigmata associated with hearing loss, or prenatal/perinatal medical history (Supplementary Statement of Joint Committee on Infant Hearing Screening July 1, 1972, 1974). However, with the growing recognition that approximately half of all children with congenital hearing loss present with no identifiable risk factors at birth, (Stein et al. 1983, 1990; Mauk et al. 1991), universal newborn hearing screening has become the standard of care in the United States and throughout much of the developed world.

Before the widespread practice of newborn hearing screening in the United States, the age at which hearing loss was identified and confirmed in young children with no recognized risk factors ranged between 13 and 24 months for those with severe and profound loss and 21 months to 4 years for those with mild to moderate loss (Harrison and Roush 1996). With the implementation of newborn hearing screening, the average age at which hearing loss is confirmed and therapeutic intervention initiated has decreased significantly (Vohr et al. 1998; Harrison et al. 2003; Sininger et al. 2009). In a recent study, Sininger et al. (2009) reported a median age at diagnosis of 3.03 months for infants who received hearing screening at birth, compared to 27.83 months for infants who were not screened. More importantly, the median ages for the fitting of amplification and initiation of early intervention services among screened infants were 5.59 and 10.58 months, respectively, compared to 31.07 and 37.33 months for nonscreened infants.

To date, most studies on outcomes of children with hearing loss identified through early identification and intervention programs have focused on language development. Although the findings from these studies may not exclusively reflect auditory-based outcomes, the major findings are worthy of mention.

4.1 Language

In one of the most widely cited studies, Yoshinaga-Itano et al. (1998) compared language abilities in 150 early- versus later-identified children with hearing loss. The children whose hearing losses were identified by 6 months of age ($n=72$) demonstrated significantly better expressive and receptive language abilities than those whose losses were identified after 6 months ($n=78$), at least through the first 3 years of life. There was no significant difference between the two age-at-identification groups on a number of other variables typically associated with language ability, including age at testing, gender, degree of hearing loss, ethnicity and socioeconomic status, communication mode (spoken language vs. combination of spoken and sign language), cognitive status, and presence of additional disabilities.

In subsequent studies, the strong association between earlier identification/intervention and better language outcomes in children with hearing loss continued to be evident out to 5 years (Moeller 2000) and even as far out as 8 years (Kennedy et al. 2006). Consistent with the findings of Yoshinaga-Itano and Apuzzo (1998) and Yoshinaga-Itano et al. (1998), language outcomes were not significantly related to

degree of hearing loss once the age at hearing loss confirmation (Kennedy et al. 2006) or enrollment in early intervention services (Moeller 2000) had been accounted for in the analysis. Earlier age at fitting of amplification (Sininger et al. 2010) or a cochlear implant (Svirsky et al. 2000; Nicholas and Geers 2006; Niparko et al. 2010) provides additional support that “earlier is better” in terms of spoken language acquisition.

The acquisition of language is not solely an auditory-based skill and the studies cited above included children using sign-supported communication as well as children communicating in the auditory–oral modality. Speech perception and production are more closely aligned to auditory abilities, and the effects of age at identification on development of these skills are summarized in the sections that follow.

4.2 Speech Perception

Few studies to date have examined the impact of early identification on speech perception skills. Recently, Sininger et al. (2010) reported on the effects of age at amplification fitting on communication outcomes in 44 children with congenital hearing losses from mild to profound in a longitudinal study on auditory development of young children with hearing loss. Age at amplification ranged from 1 to 72 months, with 23 subjects being fitted before the age of 6 months. The two speech perception measures administered were the On-line Imitative Test of Speech Pattern Contrast Perception (OLIMSPAC; Eisenberg et al. 2003; Boothroyd et al. 2010) and the Pediatric Speech Intelligibility Test (PSI; Jerger et al. 1980), both of which were developed for young children. The OLIMSPAC assesses children’s ability to identify phonologically significant contrasts by imitating VCV stimuli. The PSI is a closed-set task that assesses word and sentence identification in quiet and single-talker competition using a picture-pointing response. Both measures were administered when the children were 3 years of age and at 6-month follow-up intervals thereafter. Regression analysis was performed to determine the importance of different variables in explaining the variance in the data, including age at amplification, degree of hearing loss, parent–child interaction, language in the home, and intensity of intervention. Age at fitting of amplification was the most important factor in predicting performance for both speech perception measures—earlier fitting corresponded to improved scores.

4.3 Speech Production

Strong relationships between age of identification/intervention and speech production outcomes per se have not been demonstrated at the time of this writing. Thus, the impact of early identification and intervention on developing speech production skills appears to be complex. Studies involving early identified children who receive

appropriate sensory intervention and habilitative services within the first year of life suggest better outcomes compared to those reported in studies conducted before the advent of universal newborn hearing screening. However, delays in speech production continue to be evident in early identified children and appear to be mediated by a combination of variables, including but not limited to the child's residual hearing, the amount of auditory access afforded by the sensory device and consistency of its use, the quality and intensity of intervention services in place, and any language advantage made possible through earlier auditory access and habilitation (Fitzpatrick et al. 2007; Moeller et al. 2007a; Yoshinaga-Itano and Sedey 2000). On a more positive note, phonologic delays have been less pronounced (Moeller et al. 2007a) and speech intelligibility ratings have been substantially higher (Yoshinaga-Itano 2000; Yoshinaga-Itano and Sedey 2000) in early-identified children when compared to results reported by earlier investigators for later-identified children.

The effect of early fitting of a sensory device on speech production development has been documented during the past decade. After cochlear implantation, onset of babbling and accelerated expansion of phonetic inventories have been demonstrated, providing compelling evidence regarding the link between early auditory experience and subsequent vocal development (Ertmer and Mellon 2001; Kishon-Rabin et al. 2005; Moeller et al. 2007a). Sininger et al. (2010) reported on the effects of age of amplification fitting on articulation development. Similar to their findings on speech perception, age at amplification fitting was a significant predictor of speech production outcomes, but was not as strongly predictive as degree of hearing loss.

Age at identification of hearing loss has been shown to be a strong predictor of successful language and speech perception development, with mixed results for speech production. This latter finding is somewhat surprising because of the strong link between perception and production, but may reflect a lag in oral-motor development.

5 Mode of Communication/Intervention

A factor that indirectly influences auditory development in children with hearing loss is mode of communication, which often intertwines with the intervention program selected for the child. Most parents of children with hearing loss entertain hopes that their child will learn to communicate via spoken language. Despite such aspirations, there is the likelihood that, with increasing severity of hearing loss, a child will need to rely on some form of manual communication (Robbins et al. 1999). Before the availability of cochlear implants, the majority of children with severe to profound hearing loss would have been placed in programs that advocated some form of sign language (Conrad 1979). Other alternatives might have included a combination of sign and oral communication termed total communication (TC), or a method that supplements lipreading called cued speech. It is notable that some children with severe to profound hearing loss have been able to succeed in programs advocating strong auditory-oral (AO) approaches with their hearing aids (Musselman and Kircaali-Iftar 1996). After cochlear implantation became standard clinical

practice, many more children with severe to profound hearing loss have had the opportunity to develop spoken language through training programs that emphasize auditory learning (Svirsky et al. 2000).

There are a number of investigators who have compared speech perception data and related skill outcomes for children from both AO and TC backgrounds (Somers 1991; Geers et al. 2000; Tobey et al. 2004). Although the results from these studies generally indicate that children in AO programs have higher scores on measures of speech perception compared to children in TC programs, it is difficult to draw clear conclusions regarding auditory development from these studies because children were not randomly assigned to a communication mode. Children in programs focused on AO development often differ significantly from children in TC programs. For example, Robbins et al. (1999) found that the children who were in TC programs had more profound hearing impairments than the children in AO programs. These investigators speculated that reduced auditory potential with hearing aids necessitated a more manually based approach to language. In addition, TC programs do not equally emphasize sign and speech. Children enrolled in TC programs that focus on auditory skill development may perform better on measures of speech perception than children in TC programs that provide with less emphasis on audition (Geers 2002). Further, the mode of communication used by a child is not necessarily static. Children often switch from AO to TC placements and vice versa at various points during their childhood. These changes in programming are often based on a child's progress. For example, children who are developing strong AO skills while in a TC program may change to an AO placement. The opposite may also occur; children who are struggling with AO development may change to a TC program. As a result of these types of changes, these groups of children are often "self-selected" by the very skill that investigators purport to examine as an outcome factor. Lastly, many children in TC programs may not have the auditory skill to be assessed on the more difficult measures of speech perception, such as open-set tests in quiet or in noise, thereby reducing the number of children in TC programs who might be able to participate in such studies (e.g., Eisenberg et al. 2004).

Despite the various factors that can complicate interpretation of results from studies that compare outcomes based on the different communication modes, Pisoni and Geers (2000) have noted that children in AO programs are trained to process speech via the auditory modality. For that reason, children who communicate using spoken language have a greater opportunity to establish robust phonological representations of the speech sounds in their language than do children trained using manually based communication modes.

6 Effects of Auditory Deprivation and Experience

There is mounting evidence that auditory deprivation due to sensorineural hearing loss in the early months of life has far-reaching consequences for the development and function of auditory pathways beyond the cochlea and auditory nerve.

Histological studies that trace emerging auditory brain stem and cortical function have identified distinct periods of maturation from birth to late childhood (Moore and Linthicum 2007; Moore and Eggermont, Chap. 3). On the basis of this development, inconsistent or absence of auditory stimulation during any of these periods is likely to affect orderly development. The impact of congenital deafness followed by partial restoration of hearing provides a rare opportunity to learn about the developing auditory system after a period of deprivation and to ascertain whether it matures in an orderly sequence after introduction of sound via sensory assistance. Select studies cited in the text that follows address the effects of auditory deprivation and experience through physiological (auditory evoked potentials) and behavioral (speech perception) outcomes.

6.1 Auditory Evoked Potentials

Evidence of restoration and maturation has been shown for at least some aspects of higher order auditory function after cochlear implantation in children with profound hearing loss. In particular, investigators have focused on P1, a response thought to emanate from auditory thalamic and cortical sources (Erwin and Buchwald 1987; McGee, and Kraus 1996). Reduction in peak latency of the P1 cortical evoked response has been designated as a marker indirectly reflecting central auditory plasticity and maturation after cochlear implantation (e.g., Ponton et al. 1996; Sharma et al. 2002).

In an early study on cortical responses in children with cochlear implants, Ponton et al. (1996) recorded P1 peak latencies in four participant groups: postlingually deafened adults and prelingually deafened children with cochlear implants, and adults and children with normal hearing. The response was elicited by clicks in the participants with normal hearing, whereas biphasic current pulses were delivered directly to the implant in the participants with hearing loss. For the children with cochlear implants, the time between age at identification of deafness and cochlear implantation ranged from 5 to 9 years. Results demonstrated that for each year of auditory deprivation, maturation of P1 latencies was delayed by 1 year. Maturation of P1 latency after cochlear implantation paralleled the normal progression of maturation. Extrapolation of the data using best-fit functions suggested that the P1 latencies of children with normal hearing would reach maturity by the age of 15 years. The P1 latencies of children with implants would reach maturity between 17 and 25.5 years of age, depending on the length of auditory deprivation.

Sharma et al. (2002) asked whether there was a sensitive period for central auditory development in children receiving a cochlear implant. P1 latency was elicited in response to the synthesized syllable /ba/ in 104 children with implants who were between the ages of 2.3 and 18 years. Responses were recorded at around 6 months post device activation. Referenced to P1 latencies of children with normal hearing, children implanted at 3 years, 6 months of age or younger demonstrated

age-appropriate P1 latencies by 6 months postactivation of the implant. Conversely, most of the children implanted between 3.5 and 7 years and all children implanted after 7 years did not show normal P1 latencies. The investigators concluded that longer periods of auditory deprivation adversely affect the development of synaptic efficiency and transmission time. For the minority of children in the middle group who did in fact produce normal P1 latencies after implantation, it was surmised that preimplant residual hearing, experience with hearing aids, and quality of habilitation may have contributed to central auditory development. Longitudinal assessment indicated little change in P1 latency after 2 years of device use in children implanted at 11 years of age or older. In children implanted by the age of 7 years, P1 latency was shown to decrease within a year of implant use. These findings, in conjunction with those of Ponton et al. (1996), are significant because they demonstrate that early auditory stimulation can negate the deleterious effects of sound deprivation in the congenitally deaf ear.

With the widespread practice of bilateral cochlear implantation, researchers have begun to explore binaural effects through electrophysiological recordings at different sites of the auditory system. In a study designed to investigate interaural timing at the level of the brainstem, Gordon et al. (2008) measured the electrical auditory brainstem response over a 9-month period to examine latency differences between implanting first and second ears in children implanted either sequentially or simultaneously. The study sample consisted of children who received their first implant before the age of 3 years. Both short (less than 1 year) and long (greater than 2 years) delays between first and second ears resulted in prolonged latencies for wave eV and interwave eIII–eV in the second implanted ear of children who underwent sequential implantation. The latency differences were twice as long for those in the long-delay group relative to the short-delay group. The latencies decreased systematically over 9 months for both groups, with significant prolongation remaining for those children in the long-delay group. No latency delays were observed for the children who underwent simultaneous implantation.

Auditory development with bilateral cochlear implantation was investigated by Bauer et al. (2006) in a retrospective analysis of P1 latency measured in response to the synthesized syllable /ba/. Response latencies were analyzed from two children with sequential implantation and two children with simultaneous implantation; their ages ranged from 12 to 24 months and the second device was implanted by 24 months of age for all four children. The P1 latencies for the children with simultaneous bilateral cochlear implantation were within normal limits by 1 month postactivation. The two children sequentially implanted received their second devices at 5 and 12 months after the first implant. The P1 latencies for the first implant reached normal limits by 3–6 months postactivation; the P1 latency for the second implant reached normal limits by 1 month postactivation. One of the most interesting findings from this study pertains to the rapid rate of maturation in simultaneously implanted children. Bauer et al. posited that a synergistic effect may occur when both crossed and uncrossed fibers of the auditory nervous system receive simultaneous stimulation.

6.2 *Speech Perception*

Many clinical investigators have studied the effects of auditory experience on speech perception outcomes in children with cochlear implants. Results from longitudinal studies have been consistent in demonstrating that open-set word scores improve steadily (Fryauf-Bertschy et al. 1997; Papsin et al. 2000; Waltzman et al. 2002), even after 10 years of cochlear implant use (Uziel et al. 2007; Geers et al. 2008). Harrison et al. (2005) asked whether there is a critical period in the acquisition of speech perception after cochlear implantation. These investigators addressed this question statistically by subjecting their clinical speech recognition to binary partitioning analysis. The data consisted of scores from different test measures administered to children implanted between 1 and 15 years of age. Although the analysis revealed “breaks in the data” at different ages for different test measures, there was no clear evidence of a single age (or length of deprivation) to suggest a critical time window for implantation.

Documenting the emerging speech perception capabilities of hearing-impaired infants is a relatively new area of investigation and primarily involves those babies who have received cochlear implants. In one of the first studies, Houston et al. (2003b) applied the Split-Screen Preferential Looking Paradigm to determine when young children with cochlear implants begin to associate speech sounds presented simultaneously with objects occurring in motion. The auditory and visual stimuli consisted of: (1) the phoneme /a/ paired with a video clip of a moving toy airplane; (2) the words /hap hap hap/ paired with a video clip of a toy kangaroo hopping; and (3) the phoneme /i/ with rising or falling intonation paired with a video clip of a rising bubble or a ball rolling down, respectively. Relative to performance of infants with normal hearing (6–30 months of age), children receiving cochlear implants between 7 and 15 months of age learned to pair auditory and visual stimuli accurately after 2–6 months of experience with the implant. Infants implanted at older ages (16–25 months) did not learn to associate these pairings even by 18 months postimplantation. The testing paradigm was designed to assess pre-word learning skills. The investigators conjectured that age at implantation in combination with length of auditory deprivation may disrupt “intersensory redundancy,” potentially impairing later word learning skills in children implanted at older ages.

The VRASPAC was adopted by Uhler et al. (2011) in a longitudinal study of speech perception in two infants with cochlear implants and two age-matched normal-hearing infants. Recall that VRASPAC is a test developed for infants that requires a conditioned head-turn response when detecting a change between contrasting stimuli (Eisenberg et al. 2007; Martinez et al. 2008). The following contrasts were used by Uhler in her study: /a-i/, /u-i/, /a-u/, /ta-da/, and /pa-ka/. The normal-hearing infants attained criteria for all five contrasts within two test sessions. The infants with cochlear implants met criteria for three of the five contrasts after 2 months of experience with the implant; by 3 months one child discriminated all five contrasts and the other child discriminated four contrasts. It is impressive that these data are fairly consistent with the Sharma et al. (2002) P1 latency results in terms of the time course of maturation.

In a study on bimodal hearing, Holt et al. (2005) investigated the time course required for children to adapt to the cochlear implant in the poorer-hearing ear while continuing to use a hearing aid in the better hearing ear. Results from this prospective investigation confirmed that bimodal stimulation improved word recognition scores both in quiet and in noise when compared to performance with either device alone. However, an average of 2 years with bimodal listening was required before these results were attained.

Studies on sequential cochlear implantation also provide an opportunity to examine the time course of bilateral processing. In one particularly comprehensive investigation, Peters et al. (2007) examined age and experience effects with a group of 30 children implanted with their first device before the age of 5 years who subsequently received a second cochlear implant in the ear opposite to the first implant. The children were divided into three groups as a function of age at time of second implant: 3–5 years, 5.1–8 years, and 8.1–13 years. Data were collected at 3, 6, and 12 months after activation of the second implant. Open-set speech recognition for the second ear emerged by 6 months. The children in the youngest age group progressed at a more rapid pace and attained significantly higher scores than the older group by the 12-month test interval. Testing in background noise demonstrated binaural enhancement by 9 months' experience with bilateral implantation when the scores were averaged across all subjects.

The studies represented here, both physiological and behavioral, focus almost exclusively on cochlear implants. Implant technology has made it possible to address questions of auditory deprivation and plasticity because the children born with profound hearing loss are essentially deprived of sound until intervention with an implant. Similar studies conducted on early identified infants at the time of hearing aid fitting would provide even greater insights into the developing auditory system of child born with hearing loss.

7 Conclusion

Auditory development in children born with permanent hearing loss is dependent on many factors, the most consequential being degree of hearing loss and age at identification/intervention. Understanding auditory development in these children is a work in progress. The classic studies were published before the widespread practice of newborn hearing screening and before the commercial availability of cochlear implants. A strong association between degree of hearing loss and auditory skill development was the conventional wisdom for many years. Contemporary research has shown these predictions to be less obvious because of the widespread use of cochlear implantation. In addition, the duration of auditory deprivation has decreased for most children born with hearing loss because of universal newborn hearing screening. Results from recent studies reflect this change in terms of improved speech perception/production and language in comparison to children exposed to much longer periods of auditory deprivation.

Evidence from different sources points to the importance of early identification in reducing the deprivation effect and promoting auditory-based outcomes in children with hearing loss. Despite these positive effects, the child born with hearing loss is still left with an imperfect auditory system. As pointed out by Moore and Shannon (2009), the acoustic stimulus being processed through a sensory device and then delivered to the central auditory system of an individual with hearing loss remains degraded. Nonetheless, children with atypical hearing from early onset have the capability to develop impressive spoken language skills through processes of physiological and perceptual reorganization at the level of the central auditory nervous system.

Auditory experience after sensory device activation affords researchers the opportunity to explore the effects of auditory deprivation and the time course of maturation through use of behavioral and electrophysiological measures. The majority of these studies have been conducted on children with cochlear implants. The most compelling evidence comes from research on auditory evoked potentials. Findings from these studies demonstrate the critical role that early auditory stimulation plays in promoting maturation of central auditory pathways and pinpointing a critical period when the auditory system is most plastic.

Investigations into the development of speech perception skills do not present a clear or consistent picture to explain auditory perceptual development in infants and toddlers with hearing loss. Few researchers have taken the initiative to track outcomes in infants at the time hearing loss is identified and hearing aids are fitted. Variations in protocols and children's ages at time of testing contribute in part to diverse outcomes. Moreover, some children transition from hearing aids to cochlear implants during the course of a study. Despite these complexities, a few researchers are beginning to track the development of speech perception in newborns identified with hearing loss. Future longitudinal studies that combine behavioral and electrophysiological measures with large samples of newly identified infants should provide a more comprehensive understanding of atypical auditory development.

Acknowledgments Preparation of this chapter was supported by the following grants from the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health: R01DC006238, R01DC004797, R01DC008875, and R01DC009561.

References

- Barker, B. A., & Tomblin, J. B. (2004). Bimodal speech perception in infant hearing aid and cochlear implant users. *Archives of Otolaryngology—Head and Neck Surgery*, 130, 582–586.
- Bauer, P. W., Sharma, A., Martine, K., & Dorman, M. (2006). Central auditory development in children with bilateral cochlear implants. *Archives of Otolaryngology—Head and Neck Surgery*, 132, 1133–1136.
- Bergeson, T. R., Pisoni, D. B., & Davis, R. A. O. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing*, 26, 149–164.

- Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, 44, 264–285.
- Boothroyd, A. (1978). Speech perception and sensorineural hearing loss. In M. Ross & T. G. Giolas (Eds.), *Auditory management of hearing-impaired children: Principles and prerequisites for intervention* (pp. 117–144). Baltimore: University Park Press.
- Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research*, 27, 134–144.
- Boothroyd, A. (1997). Auditory development of the hearing child. *Scandinavian Audiology*, 26(Supplement 46), 9–16.
- Boothroyd, A., & Cawkwell, S. (1970). Vibrotactile thresholds in pure-tone audiometry. *Acta Oto-Laryngologica*, 69, 384–387.
- Boothroyd, A., & Medwetsky, L. (1992). Spectral distribution of /s/ and the frequency response of hearing aids. *Ear and Hearing*, 13, 150–157.
- Boothroyd, A., Eisenberg, L. S., & Martinez, A. S. (2010). An on-line imitative test of speech pattern contrast perception (OlimSpac): Developmental effects in normally hearing children. *Journal of Speech, Language, and Hearing Research*, 53, 531–542.
- Ching, T. Y. C., Psarros, C., Hill, M., Dillon, H., & Incerti, P. (2001). Should children who use cochlear implants wear hearing aids in the opposite ear? *Ear and Hearing*, 22, 365–380.
- Ching, T. Y. C., van Wanrooy, E., & Dillon, H. (2007). Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: A review. *Trends in Amplification*, 14, 161–192.
- Conrad, R. (1979). *The deaf school child*. London: Harper & Row.
- Davis, B. L., Morrison, H. M., Von Hapsberg, D., & Warner-Czyz, A. (2005). Early vocal patterns in infants with varied hearing levels. *Volta Review*, 105, 7–27.
- Davis, H. (1965). The young deaf child; Identification and management. *Acta Oto-Laryngologica* 206(Supplement), 1–258.
- Davis, J. M., Elfenbein, J., Schum, R., & Bentler, R. A. (1986). Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children. *Journal of Speech and Hearing Disorders*, 51, 53–62.
- Downs, M. P. (1970). The identification of congenital deafness. *Transactions of the American Academy of Ophthalmology and Otorhinolaryngology*, 74, 1208–1214.
- Eisenberg, L. S., Martinez, A. S., & Boothroyd, A. (2003). Auditory-visual and auditory-only perception of phonetic contrasts in children. *Volta Review*, 103, 327–346.
- Eisenberg, L. S., Kirk, K. I., Martinez, A. S., Ying, E. A., & Miyamoto, R. T. (2004). Communication abilities of children with aided residual hearing: Comparison with cochlear implant users. *Archives of Otolaryngology—Head and Neck Surgery*, 130, 563–569.
- Eisenberg, L.S., Martinez, A.S., & Boothroyd, A. (2007). Assessing auditory capabilities in young children. *International Journal of Pediatric Otorhinolaryngology*, 71, 1339–1350.
- Elfenbein, J. L., Hardin-Jones, M. A., & Davis, J. M. (1994). Oral communication skills of children who are hard of hearing. *Journal of Speech and Hearing Research*, 37, 216–226.
- Elliott, L., & Katz, D. (1980). *Development of a new children's test of speech discrimination*. St. Louis: Auditec.
- Erber, N. P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech and Hearing Research*, 15, 413–422.
- Ertmer, D. J., & Mellon, J. A. (2001). Beginning to talk at 20 months: Early vocal development in a young cochlear implant recipient. *Journal of Speech, Language, and Hearing Research*, 44, 192–206.
- Ertmer, D. J., Young, N., & Nathani, S. (2007). Profiles of vocal development in young cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, 50, 393–407.
- Erwin, R. J., & Buchwald, J. S. (1987). Midlatency auditory evoked responses in the human and the cat model. *Electroencephalography and Clinical Neurophysiology Supplement*, 40, 461–567.

- Fitzpatrick, E., Durieux-Smith, A., Eriks-Brophy, A., Olds, J., & Gaines, R. (2007). The impact of newborn hearing screening on communication development. *Journal of Medical Screening*, 14, 123–131.
- Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D. M. T., Gantz, B. J., & Woodworth, G. G. (1997). Cochlear implant use by prelingually deafened children: The influence of age at implant and length of device use. *Journal of Speech, Language, and Hearing Research*, 40, 183–199.
- Geers, A. E. (2002). Factors affecting the development of speech, language, and literacy in children with early cochlear implantation. *Language, Speech, and Hearing Services in Schools*, 33, 172–183.
- Geers, A. E., Nicholas, J., Tye-Murray, N., Uchanski, R., Brenner, C., Davidson, L. S., Toretta, G., & Tobey, E. A. (2000). Effects of communication mode of long-term cochlear implant users. *Annals of Otolaryngology, Rhinology, and Laryngology*, 109(Supplement 185), 89–92.
- Geers, A., Tobey, E., Moog, J., & Brenner, C. (2008). Long-term outcomes of cochlear implantation in the preschool years. *International Journal of Audiology*, 47(Supplement 2), S21–S30.
- Gordon, K. A., Twitchell, K. A., Papsin, B. C., & Harrison, R. V. (2001). Effect of residual hearing prior to cochlear implantation on speech perception in children. *Journal of Otolaryngology*, 30, 216–223.
- Gordon, K. A., Valero, J., van Hoesel, R., & Papsin, B. C. (2008). Abnormal timing delays in auditory brainstem responses evoked by bilateral cochlear implant use in children. *Otology & Neurotology*, 29, 193–198.
- Hack, Z. C., & Erber, N. P. (1982). Auditory, visual, and auditory-visual perception of vowels by hearing-impaired children. *Journal of Speech and Hearing Research*, 25, 100–107.
- Harrison, M., & Roush, J. (1996). Age of suspicion, identification and intervention for infants and young children with hearing loss: A national study. *Ear and Hearing*, 17, 55–62.
- Harrison, M., Roush, J., & Wallace, J. (2003). Trends in age of identification and intervention in infants with hearing loss. *Ear and Hearing*, 24, 89–95.
- Harrison, R. V., Gordon, K. A., & Mount, R. J. (2005). Is there a critical period for cochlear implantation in congenitally deaf children? Analysis of hearing and speech perception performance after implantation. *Developmental Psychobiology*, 46, 252–261.
- Haskins, H. (1949). A phonetically balanced test of speech discrimination for children. Master's thesis, Northwestern University, Evanston, IL.
- Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, 45, 573–584.
- Holt R. F., Kirk, K. I., Eisenberg, L. S., Martinez, A. S., & Campbell, W. (2005). Spoken word recognition development in children with residual hearing using cochlear implants and hearing aids in opposite ears. *Ear and Hearing*, 26(4 Supplement), 82 S–91 S.
- Houston, D. (2002). What infants learn about native language sounds organization during their first year, and what may happen if they don't. In K. Schauwers, P. J Govaerts, & S. Gillis (Eds.), *Language acquisition in very young children with a cochlear implant*. Antwerp Papers in Linguistics (102). Retrieved July 9, 2009, from <http://webh01.ua.ac.be/apil/apil102/apil102.pdf>.
- Houston, D. M., Pisoni, D. B., Kirk, K. I., Ying, E. A., & Miyamoto, R. T. (2003a). Speech perception skills of deaf infants following cochlear implantation: A first report. *International Journal of Pediatric Otorhinolaryngology*, 67, 479–495.
- Houston, D. M., Ying, E. A., Pisoni, D. B., & Kirk, K. I. (2003b). Development of pre-word learning skills in infants with cochlear implants. *Volta Review*, 103, 303–326.
- Jerger, S., Lewis, S., Hawkins, J., & Jerger, J. (1980). Pediatric speech intelligibility test. I. Generation of test materials. *International Journal of Pediatric Otorhinolaryngology*, 2, 217–230.
- Kennedy, C., McCann, D., Campbell, M.J., Law, C., Mullee, M., Petrou, S., Watkin, P., Worsfold, S., Yuen, H. M., & Stevenson, J. (2006). Language ability after early detection of permanent childhood hearing impairment. *The New England Journal of Medicine*, 354, 2131–2141.
- Kishon-Rabin, L., Taitelbaum-Swead, R., Ezrati-Vinacour, R., & Hildesheimer, M. (2005). Prelexical vocalization in normal hearing and hearing-impaired infants before and after cochlear

- implantation and its relation to early auditory skills. *Ear and Hearing*, 26(4 Supplement), 17S-29S.
- Lachs, L., Pisoni, D. B., & Kirk, K.I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear and Hearing*, 22, 236-251.
- Levitt, H., McGarr, N. S., & Geffner, D. (Eds.). (1987). Development of language and communication skills in hearing-impaired children. *ASHA Monographs*, Number 26. Rockville Pike: American Speech-Language-Hearing Association.
- Litovsky, R. Y., Johnstone, P. M., & Godar, S. P. (2006). Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology*, 45(Supplement 1), 78-91.
- Mace, A. L., Wallace, K. L., Whan, M. Q., & Stelmachowicz, P. G. (1991). Relevant factors in the identification of hearing loss. *Ear and Hearing*, 12, 287-293.
- Martinez, A., Eisenberg, L., Boothroyd, A., & Visser-Dumont, L. (2008). Assessing speech pattern contrast perception in infants: Early results on VRASPAC. *Otology & Neurotology*, 29, 183-188.
- Mauk, G. W., White, K. R., Mortenson, L. B., & Behrens, T. R. (1991). The effectiveness of screening programs based on high risk characteristics in early identification of hearing impairment. *Ear and Hearing*, 12, 312-319.
- McGee, T., & Kraus, N. (1996). Auditory development reflected by middle latency response. *Ear and Hearing*, 17, 419-429.
- Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, 106, 1-9.
- Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., Wood, S., Lewis, D., Pittman, A., & Stelmachowicz, P. (2007a). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I - Phonetic development. *Ear and Hearing*, 28, 605-627.
- Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., Wood, S., Lewis, D., Estee, S., Pittman, A., & Stelmachowicz, P. (2007b). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part II - Transition to words. *Ear and Hearing*, 28, 628-642.
- Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing*, 17, 133-161.
- Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: A timeline of development. *International Journal of Audiology*, 46, 460-478.
- Moore, D. R., & Shannon, R. V. (2009). Beyond cochlear implants: Awaking the deafened brain. *Nature Neuroscience*, 12, 686-691.
- Musselman, C., & Kircaali-Iftar, G. (1996). The development of spoken language in deaf children: Explaining the unexplained variance. *Journal of Deaf Studies and Deaf Education*, 1(2), 109-121.
- Nicholas, J. G., & Geers, A. E. (2006). Effects of early auditory experience on the spoken language of deaf children at 3 years of age. *Ear and Hearing*, 27, 286-298.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., & Fink, N. E., for the, CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association*, 303, 1498-1506.
- Nott, P., Cowan, R., Brown, M., & Wigglesworth, G. (2009). Early language development in children with profound hearing loss fitted with a device at a young age. Part I The time taken to acquire first words and first word combinations. *Ear and Hearing*, 30, 526-540.
- Oller, D. K., Wieman, L. A., Doyle, W., & Ross, C. (1975). Infant babbling and speech. *Journal of Child Language*, 3, 1-11.
- Osberger, M. J. (Ed.) (1986). Language and learning skills of hearing-impaired students. *ASHA Monographs*. Rockville Pike, MD: American Speech-Language-Hearing Association.
- Papsin, B. C., Gysin, C., Picton, N., Nedzelski, J., & Harrison, R. V. (2000). Speech perception outcome measures in prelingually deaf children up to four years after cochlear implantation. *Annals of Otolaryngology, Rhinology, and Laryngology*, 109(Supplement 185), 38-42.

- Peters, B. R., Litovsky, R., Parkinson, A., & Lake, J. (2007). Importance of age and postimplantation experience on speech perception measures in children with sequential bilateral cochlear implants. *Otology & Neurotology*, 28, 649–657.
- Pisoni, D. B., & Geers, A. E. (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing. *Annals of Otology, Rhinology and Laryngology*, 109(Supplement 185), 92–93.
- Ponton, C. M., Don, M., Eggermont, J. J., Waring, M. D., & Masuda, A. (1996). Auditory system plasticity in children after long periods of complete deafness. *NeuroReport*, 8, 61–65.
- Robbins, A. M., Bollard, P. M., & Green, J. (1999). Language development in children implanted with the Clarion® Cochlear Implant. *Annals of Otology, Rhinology, and Laryngology*, 108(4), 113–118.
- Schauwers, K., Gillis, S., Daemers, K., De Beukelaer, C., & Govaerts, P. J. (2004). Cochlear implantation between 5 and 20 months of age: The onset of babbling and the audiologic outcome. *Otology & Neurotology*, 25, 263–270.
- Scollie, S., Glista, D., Bagatto, M., & Seewald, R. (2007). Multichannel nonlinear frequency compression: A new technology for children with hearing loss. In R. C. Seewald & J. M. Bamford (Eds.), *A sound foundation through early amplification 2007: Proceedings of the Fourth International Conference* (pp. 151–159). Chicago: Phonak AG.
- Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, 23, 532–539.
- Sininger, Y. S., Martinez, A., Eisenberg, L., Christensen, E., Grimes, A., & Hu, J. (2009). Newborn hearing screening speeds diagnosis and access to intervention by 20–25 months. *Journal of the American Academy of Audiology*, 20, 49–57.
- Sininger, Y. S., Grimes, A., & Christensen, E. (2010). Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss. *Ear and Hearing*, 31, 166–185.
- Somers, M. N. (1991). Speech perception abilities in children with cochlear implants or hearing aids. *The American Journal of Otology*, 12(Supplement), 174–178.
- Stein, L., Clark, S., & Kraus, N. (1983). The hearing-impaired infant: Patterns of identification and rehabilitation. *Ear and Hearing*, 4, 232–236.
- Stein, L., Jabaley, T., Spitz, R., Stoakley, D., & McGee, T. (1990). The hearing-impaired infant: Patterns of identification and rehabilitation revisited. *Ear and Hearing*, 11, 201–205.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., & Lewis, D. E. (2001). Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *Journal of the Acoustical Society of America*, 110, 2183–2190.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., & Lewis, D. E. (2002). Aided perception of /s/ and /z/ by hearing-impaired children. *Ear and Hearing*, 23, 316–324.
- Straatman, L. V., Rietveld, A. C., Beijen, J., Mylanus, E. A., & Mens, L. H. (2010). Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. *Journal of the Acoustical Society of America*, 128, 1885–1895.
- Supplementary Statement of Joint Committee on Infant Hearing Screening, July 1, 1972. (1974). *ASHA*, 16, 160.
- Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B., & Miyamoto, R. T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science*, 22(2), 153–158.
- Tobey, E. A., Rekart, D., Buckley, K., & Geers, A. E. (2004). Mode of communication and classroom placement impact on speech intelligibility. *Archives of Otolaryngology-Head and Neck Surgery*, 130(5), 639–643.
- Uhler, K., Yoshinaga-Itano, C., Gabbard, S. A., Rothpletz, A. M., & Jenkins, H. (2011). Longitudinal infant speech perception in young cochlear implant users. *Journal of the American Academy of Audiology*, 22, 129–142.

- Uziel, A. S., Sillon, M., Vieu, A., Artieres, F., Prion, J-P., Daures, J. P., & Mondain, M. (2007). Ten-year follow-up of a consecutive series of children with multichannel cochlear implants. *Otology & Neurotology*, 28, 615–628.
- Vihman, M., Ferguson, C. A., & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual difference. *Applied Psycholinguistics*, 7, 3–40.
- Vohr, B. H., Carty, L. M., Moore, P. E., & Letourneau, K. (1998). The Rhode Island hearing assessment program: Experience with statewide screening (1993–1996). *Journal of Pediatrics*, 133, 353–357.
- Wake, M., Hughes, E. K., Poulakis, Z., Collins, C., & Rickards, F. W. (2004). Outcomes of children with mild-profound congenital hearing loss at 7 to 8 years: A population study. *Ear and Hearing*, 25, 1–8.
- Waltzman, S. B., Cohen, N. L., Green J., & Roland, J. T., Jr. (2002). Long-term effects of cochlear implants in children. *Otolaryngology—Head and Neck Surgery*, 125, 505–111.
- Worthington, D. W., Stelmachowicz, P., & Larson, I. (1986). Audiological evaluation. In M. J. Osberger (Ed.), *Language and learning skills in hearing-impaired students* (pp. 12–20). ASHA Monographs, 23. Rockville, MD: American Speech-Language-Hearing Association.
- Yoshinaga-Itano, C. (2000). Development of audition and speech: Implications for early intervention with infants who are deaf or hard of hearing. *Volta Review*, 100, 213–234.
- Yoshinaga-Itano, C., & Apuzzo, M. L. (1998). Identification of hearing loss after age 18 months is not early enough. *American Annals of the Deaf*, 143, 380–387.
- Yoshinaga-Itano, C., & Sedey, A. (2000). Early speech development in children who are deaf or hard of hearing: Interrelationships with language and hearing. *Volta Review*, 100, 181–211.
- Yoshinaga-Itano, C., Sedey, A., Coulter, D., & Mehl, A. (1998). Language of early and later-identified children with hearing loss. *Pediatrics*, 102, 1161–1171.

Index

A

- Acoustic cues, stream segregation, 140–141
- Acoustic reflex, development, 27–28
- Across-channel processing, psychophysics, 123ff
- Active processes, cochlea, 42ff
- Adults, auditory deprivation, 187–188
 - auditory stream segregation compared to infants, 140
 - binaural hearing, 165–167
 - binaural unmasking, 166–167
 - melody recognition, 236–237
 - musicians vs. nonmusicians, 241ff
 - sound localization, 165
 - sound source discrimination, 165
- Adults, speech perception compared to infants, 213–215
- Age-related changes, 2
 - auditory scene analysis, 137ff
- Amplification, newborns, 43–44
- Analysis system, maturation, 94
- Attention development, in psychophysics, 125–126
- Attention system, maturation, 94–95
- Attention, 7ff
 - auditory, 152–153
- Attentional pruning of language, infants, 205ff
- Auditory attention, development, 137ff
 - selective, 152ff
 - types, 152–153
- Auditory brainstem response (ABR), 5, 11–12
 - effects of deprivation, 88ff
 - maturation, 67ff
 - See also Auditory Evoked Potentials
 - See structure-function relationships, 76ff
- Auditory coding, psychophysics, 107ff
- Auditory cortex, 3
 - Brodman areas, 80ff
 - translational invariance, 237
- Auditory cues, binaural hearing, 164
- Auditory deprivation, 12
 - adults, 187–188
 - effects on auditory pathway development, 267ff
- Auditory development, atypical, 255ff
 - effects of age identification of hearing loss, 263ff
 - effects of cochlear implants, 267ff
 - effects of communication intervention, 266–267
 - effects of sign language, 266–267
 - stages of, 6–7
- Auditory discrimination, and cortical maturation, 62–63
- Auditory evoked potentials, *see also* Auditory Brainstem Response
 - development of auditory pathway, 268–269
 - development, 61ff
- Auditory experience, speech perception, 270–271
- Auditory nerve, development, 29
- Auditory nervous system, development, 61ff
- Auditory pathway, effects of cochlear implants, 268–269
 - effects of experience on development, 267ff
 - effects on development of auditory deprivation, 267ff
- Auditory perception, and cortical maturation, 63
- Auditory scene analysis, 7ff
 - development, 137ff
- Auditory sensitivity in quiet, 111

- Auditory stream segregation, 139–140
 acoustic cues, 140–141
 adults vs. infants, 140
 infants, 140–141
 preschoolers and school-aged children, 141–142
- Auditory-oral approaches, speech development, 266–267
- B**
- Basic pitch perception, development, 228ff
- Basilar membrane motion, DPOAE, 34ff
- Bilateral cochlear implants, children, 184–186, 269
 sound localization, 186–187
 speech intelligibility in children, 188–190
- Bilingual infants, 206–207
- Binaural hearing, adults, 165–167
 auditory cues, 164
 children with cochlear implants 183ff
 development, 163ff
 hearing impairment, 183ff
- Binaural intelligibility level difference, 166–167
- Binaural sensitivity, development, 178–179
- Binaural unmasking, adults, 166–167
 development, 178–179
- Brain imaging, functional, 61ff
- Brain, musicians vs. nonmusicians, 241ff
- Brainstem electrophysiology, maturation, 67ff
- Brainstem response, effects of deprivation, 88ff
- Brainstem structural maturation, 65ff
- C**
- Central nervous system, musicians vs. nonmusicians, 241ff
- Children with bilateral cochlear implants, sound localization, 186–187
- Children with cochlear implants, binaural hearing, 183ff
 spatial hearing acuity, 184–186
- Children, auditory scene analysis, 137ff
 auditory stream segregation, 141–142
 bilateral cochlear implants, 184–186
 cochlear implants, 183ff, 257–258
 degrees of hearing loss, 258ff
 effect of age of identification of hearing loss, 263ff
 effects of cochlear implants on auditory development, 268–269
 hearing aids and language development, 256–258
 hearing loss and language development, 255ff
 hearing loss and sensory experience, 256–258
 informational masking, 146ff
 language development with hearing loss, 264–265
 learning harmonic syntax, 239–240
 learning music, 246–247
 minimum audible angle, 169–171
 precedence effect, 174ff
see also Infants
 sound source localization, 171–173
 spatial hearing, 169–171
 speech intelligibility, 179ff
 speech perception and auditory experience, 270–271
 speech perception with hearing impairment, 258ff, 270–271
 speech production with hearing loss, 262–263, 265–266
- Cochlea, active processes, 42ff
- Cochlear anatomy, development, 28ff
- Cochlear apex, otoacoustic emission phase, 44ff
- Cochlear frequency selectivity, development, 37ff
- Cochlear function, development, 28ff
- Cochlear implants, 12
- Cochlear implants, bilateral and speech intelligibility, 188–190
 bilateral in children, 269
 children, 183ff, 257–258
 cortical response, 92
 effects on auditory development in hearing impaired children, 268–269
 effects on auditory pathway, 268–269
 effects on speech perception in children, 270–271
 spatial hearing in children, 183–184
- Cochlear response, growth, 36
- Cocktail party problem, 153
- Comodulation detection differences, maturation, 124–125
- Comodulation masking release, maturation, 124–125
- Complex listening, children, 2–3
- Conditioned reflexes, psychophysics, 108
- Consonance/dissonance continuum, 233ff
- Consonance/dissonance perception, development, 233ff
- Consonance/dissonance, infants, 233ff
- Cortex, imaging, 82ff
 maturation (table), 78
 sink-source analysis, 86ff
 structural maturation, 79ff

Cortical activity, electrophysiology, 73ff
 Cortical function, electrophysiological maturation, 84ff
 Cortical maturation, 62
 Cortical response, early-developing, 90
 Cortical response, late-developing, 90ff

D

Degree of hearing loss, children, 258ff
 Dendritic arbors, maturation, 66
 Development, auditory attention, 137ff
 auditory scene analysis, 137ff
 basic pitch perception, 228ff
 binaural sensitivity, 163ff, 178–179
 binaural unmasking, 178–179
 human, 1ff
 informational masking, 143ff
 music perception, 223ff
 musical pitch perception, 228ff
 pitch perception, 223ff
 release from informational masking, 149–151
 spatial hearing acuity, 169–171
 spatial hearing, 163ff, 169ff
 speech perception, 197ff
 Discrimination system maturation, neural correlates, 64ff
 Discrimination system, anatomical and physiological correlations, 93–94
 maturation of ABR, 88–89
 Dissonance and consonance perception, development, 233ff
 Distortion product OAE, 31ff
 suppression, 49ff
 Distraction, in detection, 117
 Duration discrimination, psychophysics, 122

E

Ear canal absorbance, 24
 Ear canal acoustics, development, 22ff
 Ear canal transfer function, forward, 24–26
 reverse, 26
 Ear canal wall motion, infants, 27
 Ear development, functional, 19ff
 morphology, 19ff
 Encoding, 4ff
 Envelope processing, maturation, 119ff
 Event-related potentials, early maturing, 70ff
 infants, 228ff
 pitch, 228ff
 Evoked potentials, maturation (table), 83
 obligatory, 75–76

Experience, and development, 3
 and hearing development, 11ff
 effects on acquisition of key membership and harmony, 245–247
 effects on atypical auditory development, 255ff
 effects on development of auditory pathway, 267ff
 effects on melodic processing, 244–245
 effects on musical pitch processing, 241ff
 processing of single tones, 242–244
 External ear, development, 20–21

F

Feature extraction, 4ff
 Fine structure processing, maturation, 121–122
 Forced-choice procedures, psychophysics, 110–111
 Forward and reverse OAE transfer functions, 46ff
 Frequency discrimination, psychophysics, 111–112
 Frequency resolution, maturation, 115ff
 Functional development, nervous system, 61ff
 Functional imaging, cortex, 75

G

Gamma-aminobutyric acid (GABA), maturation, 70ff, 92
 Gap detection, maturation, 119–120
 Glutamate, maturation, 77–78

H

Harmonic syntax, learning by children, 239–240
 Hearing aids, language development in children, 256–258
 Hearing impairment, binaural hearing, 183ff
 Hearing loss and language development, children, 255ff
 age of identification, 263ff
 categories, 255–256
 correlation with speech perception, 258ff
 effect of age identification on auditory development, 263ff
 effects on speech, 263ff
 hearing impaired children with hearing assistance, 263ff
 language development, 255ff
 Human auditory development, 1ff
 Human development, ear, 19ff

I

- Imaging, brainstem maturation, 66–67
- Inattentiveness, 153
- Individual differences, in psychophysics, 126
- Infant auditory stream segregation, compared to adults, 140
- Infant-directed speech, 200–201
- Infants, *see also* Children
 - assessing speech perception, 260–261
 - attention to language structure, 202ff
 - attention to speech, 203ff
 - attentional pruning of language, 205ff
 - auditory experience and speech perception, 270–271
 - auditory scene analysis, 137ff
 - auditory stream segregation, 140–141
 - bilingual, 206–207
 - cochlear implants and speech perception, 270–271
 - consonance/dissonance, 233ff
 - effects of cochlear implants on auditory development, 268–269
 - event-related potentials, 228ff
 - informational masking, 146ff
 - language function, 201–202
 - learning harmonic syntax, 239–240
 - learning musical scales and tonal hierarchy, 237–239
 - masking and speech perception, 213–215
 - measuring speech perception and language skills, 198–200
 - melody recognition, 236–237
 - music, 223–224
 - perception of tonal languages, 205–206
 - perceptual attunement of language, 205ff
 - precedence effect, 174ff
 - preference for native language, 204–205
 - recognition of phonemes, 204ff
 - segmentation of speech, 208–210
 - sound localization, 173–174
 - speech perception in nonoptimal listening conditions, 213–215
 - tone discrimination, 228–229
- Informational masking, development, 143ff
 - infants and children, 146ff
 - paradigms, 143–145
 - release from 149–151
 - susceptibility, 146ff
- Inner hair cells, development, 28–29
- Intensity discrimination, psychophysics, 111ff
- Interaural level difference (ILD), binaural hearing, 164
- Interaural time difference (ITD) processing, maturation, 121–122

- Interaural time difference, binaural hearing, 164
- Internal noise, intensity discrimination, 128–129
 - masking development, 127–128

L

- Language development, children with hearing loss, 264–265
 - effects of hearing loss, 255ff
 - toddlers, 210–212
- Language function, 201–202
- Language skills, measuring in, 198–200
- Language structure, attention in, 202ff
- Language, perceptual attunement in, 205ff
 - harmonic syntax, 239–240
- Learning, musical scales and tonal hierarchy, 237–239
- Listening behavior, maturation, 127
- Loudness judgments, psychophysics, 114

M

- Maskers, speech intelligibility, 179ff
 - infant speech perception, 213–215
 - informational, 143ff
 - nonsimultaneous, 117–118
 - simultaneous, 115ff
 - spatial release from, 166–167
- Maturation, diffuse-field effects, 21–22
- Maturational changes, model, 97ff
- Medial olivocochlear (MOC) reflex, 39ff
- Melodic processing, effects of experience, 244–245
- Melody recognition, 236–237
 - adults, 236–237
- Memory development, in psychophysics, 125–126
- Middle ear transfer function, forward, 26–27
- Middle ear, development, 20–21
- Middle latency response (MLR), maturation, 68–69
- Minimum audible angle (MAA), 165
- Minimum audible angle, children, 169–171
- Mismatched negativity (MMN), maturation, 74ff 126
- Morphological development, nervous system, 61ff
- Motivation development, in psychophysics, 125–126
- Music perception, 3, 13
 - development, 223ff
 - musicians vs. nonmusicians, 241ff
- Music, in human cultures, 223–224
 - pitch perception, 224ff

- see also* Musical western, 224ff
 Musical pitch perception, development, 228ff
 Musical pitch processing, effects of experience, 241ff
 Musical pitch structure, 224ff
 Musical pitch, translational invariance, 236–237
 Musical scales, learning by children, 237–239
 Musical structure, 224ff
 Musical, *see also* Music
 Musicians vs. nonmusicians, music perception, 241ff
 Myelination, maturation, 69
- N**
- Native language, infant preference, 204–205
 Nervous system, consonance/dissonance, 235–238
 pitch encoding, 225–226
 Newborn, hearing screening, 260
 Noise, infant speech perception, 213–215
- O**
- Obligatory evoked potentials, late-maturing, 97ff
 Observer-based methods, psychophysics, 108ff
 Otoacoustic emission phase, apex of cochlea, 44ff
 Otoacoustic emissions, 5, 20
 cochlear mechanisms, 31ff, 42FF
 ear canal and middle-ear explanations, 46ff
 maturation, 30ff
 reflection in the cochlea, 42ff
 response growth, 46ff
 Outer hair cells, development, 28–29
- P**
- Pediatrics, *see* Children
 Perception, consonance/dissonance, 233ff
 Perception, physiological correlates, 95–96
 Perceptual attunement of language, 205ff
 Perceptual system maturation, neural correlates, 78ff
 Phonemes, recognition by, 204ff
 Pitch perception, development, 223ff, 228ff
 Pitch perception, music, 224ff
 Pitch structure, music, 224ff
 Pitch, encoding in nervous system, 225–226
 event-related potentials, 228ff
 Postnatal development, 2
 Precedence effect, 167
- children and, 174ff
 Psychoacoustic development, time course, 13–14
 Psychophysics, auditory coding, 107ff
 measurement techniques, 107ff
 spatial hearing, 169ff
- R**
- Reticular activating system (RAS), maturation, 70
- S**
- Segmentation of speech, 208–210
 Selective attention, 10–11, 152ff
 Sensory experience, childhood hearing loss, 256–258
 Shadowing paradigm, 153–155
 Sign language, effects on auditory development, 266–267
 Single tone processing, effects of experience, 242–244
 Sound localization, adults, 165
 children with bilateral cochlear implants, 186–187
 children, 171–173
 Sound localization, 173–174
 Sound source discrimination, adults, 165
 Source localization, behavioral measures, 167–169
 Source location discrimination, 168–169
 Spatial hearing acuity, children with cochlear implants, 184–186
 development, 169–171
 Spatial hearing children with cochlear implants, 183–184
 Spatial hearing, 3
 children, 169–171
 development, 163ff, 169ff
 psychophysical studies, 169ff
 Spatial release from masking, 166–167
 Spectral shape processing, maturation, 123–124
 Speech attention, 203ff
 Speech development, use of auditory-oral approaches, 266–267
 Speech intelligibility, children with bilateral cochlear implants, 188–190
 children, 179ff
 presence of maskers, 179ff
 Speech perception in nonoptimal listening conditions, 213–215

- Speech perception, 13, 260–261
 - adults vs. children, 213–215
 - auditory experience, 270–271
 - children with hearing impairment, 258ff
 - children with hearing loss, 265
 - correlation with hearing loss, 258ff
 - development, 197ff
 - effects of communication intervention, 266–267
 - hearing impaired children, 270–271
 - measuring in children, 198–200
 - Speech processing, 3
 - Speech production, children with hearing loss, 262–263, 265–266
 - Speech, infant-directed, 200–201
 - Spontaneous OAE, 30ff
 - immaturities, 52–53
 - Stream segregation cues, 8, 138–139
 - Suppression tuning curves, OAE, 37ff, 43–44
 - Suppression, DPOAE, 49ff
 - Susceptibility, informational masking, 146ff
- T**
- Temporal fine structure, 5
 - Temporal integration, maturation, 118–119
 - Temporal modulation transfer function (TMTF), maturation, 120
 - Temporal processing, maturation, 118ff
 - Toddlers, language development, 210–212
 - Tonal hierarchy, learning by children, 237–239
 - Tonal languages, perception by 205–206
 - Tone discrimination, infant, 228–229
 - Transfer function, middle ear, 21–22
 - Translational invariance, auditory cortex, 237
 - musical pitch, 236–237
 - Tympanic membrane, development, 20–21