Chapter 3. Development of the auditory system

1 author:

Ruth Litovsky

University of Wisconsin–Madison

162 PUBLICATIONS   **4,570** CITATIONS

SEE PROFILE

Chapter 3

Development of the auditory system

RUTH LITOVSKY*

Binaural Hearing and Speech Laboratory, Waisman Center, University of Wisconsin-Madison, Madison, WI, USA

INTRODUCTION

When considering auditory development, we may ask ourselves: what is the purpose of auditory perception and processing of auditory signals? Generally, the auditory system serves the role of constructing a perceptual space that extracts information from objects (animate and inanimate), groups together some objects, and segregates sounds from one another. In other words, the auditory system engages in an analysis of the auditory world, so that the listener can accomplish the goal of communication and learning. While the auditory system regularly interacts with other sensory modalities, much of what we understand about auditory perception and development is rooted in studies that focus the auditory signal *per se*. Auditory processing occurs in a series of analysis stages, beginning with peripheral auditory mechanisms devoted to encoding sound, and proceeding to more complex stages at which sound processing leads to perception and object recognition. An important feature of the staging is the combination of inputs from the right and left ears, which takes place several synapses after the initial impact of sound on the auditory sensory organ, the cochlea.

It is in this context that we can begin to consider the goal of auditory development, the conditions under which changes in auditory mechanisms generally result in beneficial outcomes for the listener, and, conversely, conditions under which changes bring about negative consequences. This all depends on whether these changes occur because the organism is experiencing a "normal" or "typical" trajectory of exposure to sound during development, or whether the changes are occurring due to unanticipated events, such as disease or trauma, in the auditory pathway or in other brain areas that interact with the processing of incoming signals or the production of sound by the individual.

Auditory development is a broadly defined term, which refers to the fact that perception is influenced by a combination of innate, genetically programmed changes in anatomy and physiology, combined with auditory experience. Two primary questions resonate with research to date. First, how does experience with and exposure to sound during early stages in development impact the ability to process sound? Second, how does lack of exposure to sound influence the anatomic organization and physiologic functioning of the auditory pathways? One way to address these questions, which is the approach taken in this chapter, is to consider developmental changes in the auditory pathway as one progresses from peripheral to central mechanisms. The peripheral system generally encodes stimuli based on temporal, spectral, and intensity cues. The subsequent extraction of complex combinations of those cues, assignment of auditory features to meaningful stimuli such as speech and music, and the role of non-auditory processing occur at more central levels of the auditory pathway. Finally, this chapter focuses attention on the development of the human auditory system; however, much is to be learned also from research done with animal models, which will be referred to as well. As such, from a methodologic perspective, there are a number of approaches taken towards describing and understanding human auditory development. Structural changes are generally studied using histologic approaches by reconstructing connectivity between neurons. In humans this would be done either by using neuroimaging techniques or with postpartum labeling of neural connections in cadavers. A second approach, to explore functional maturation, also relies on imaging, primarily functional

---

*Correspondence to: Ruth Litovsky, Binaural Hearing and Speech Laboratory, Waisman Center, University of Wisconsin-Madison, 1500 Highland Ave, #521, Madison WI 53705, USA. Tel: +1-608-262-5045, E-mail: litovsky@waisman.wisc.edu

neuroimaging techniques. Third, behavioral studies, which depend on extracting information from listeners who are performing psychoacoustic tasks, provide a window into the extent to which underlying structure and function support perception.

DEVELOPMENT OF THE EAR

The ear is responsible for the initial encoding of acoustic input. The functional status of the ear has been studied using various investigational and clinical tools. Anatomic studies suggest than numerous changes occur after full-term birth. For example, the anatomy of the human ear canal continues to mature; in infants the canal is shorter and straighter than it is in adults (Northern and Downs, 1984). Another example is the increase in ear canal diameter and length during the first 2 years of life (Keefe et al., 1993). Finally, even more prolonged is the development of the middle-ear cavity volume, which extends into the late teenage to years, and is likely to influence the mechanics involved in middle-ear function (Eby and Nadol, 1986). For a full review of these and other changes, see Abdala and Keefe (2011). The postnatal maturation of the external and middle ear is important to note, because they can have substantial effects on how sounds are absorbed, processed, filtered, and transmitted to the auditory system.

Cochlear development in mammals generally proceeds from the basal region to the apex of the cochlea, that is, on a gradient that follows high-to-low-frequency organization (Pujol and Hilding, 1973). The mammalian cochlea contains two types of sensory cells (the outer and inner hair cells). The innervation from the eighth nerve consists of primarily two types of ganglion neurons (type I and type II).

In utero, hair cells are fairly mature from a morphologic standpoint some time in the third trimester, although synaptic connections with the auditory nerve likely continue to mature after birth. Although not fully mature, the developing auditory system enables response to sound *in utero*, as has been shown by studies that measure fetal movement or heart rate response to sound at around 25–27 weeks' gestation. Another estimate of the onset of hearing is 27–28 weeks, obtained with the ubiquitous electrophysiologic measures of functionally, the auditory brainstem response, which shows synchronous eighth-nerve activity and average brainstem responses (Galambos and Hecox, 1978). Little is known about maturation of the basilar membrane, a structure within the inner ear that serves as the anatomic base upon which the sensory hair cells are positioned. Functionally, the basilar membrane changes in mass and elasticity along its basal-to-apical distribution, resulting in a "place–frequency" mapping resulting from resonance dependence response to sound vibrations at different frequencies.

Peripheral maturation of the auditory system can be studied using a non-invasive and clinically relevant technique whereby otoacoustic emissions (OAEs) are recorded with a small sensitive microphone inserted in the entrance to the ear canal. OAEs reflect normal physiology of the outer hair cells, which act as a cochlear amplifier. OAEs are generally present at birth; however, numerous types of OAEs have been described, which vary by the region in the cochlea and mechanisms responsible for their functionality. OAEs have become a highly utilized tool for investigating maturation of cochlear function, and thus also for identifying immature and abnormal peripheral auditory function. At birth, in contrast with the immature outer and middle ear, the inner ear seems to be more mature, as characterized by OAEs. Studies generally suggest that newborn cochleae produce robust responses that are adult-like in many ways. Nuanced research into this area further suggests that maturational changes occur particularly in the apical region of the cochlea, where low-frequency information is generally processed. How these maturational changes influence perceptual outcomes is yet to be understood (Keefe and Abdala, 2011).

BEHAVIORAL TESTING AND PSYCHOACOUSTICS

Research reviewed in this chapter focuses on how the human auditory system develops, what aspects of sound are perceived by infants, children and non-human animals, and the ways in which those sounds are interpreted at higher levels of neural encoding. It is thus prudent to begin by understanding how perception can be measured reliably. In humans, over the lifespan, depending on the age of the infant/toddler/child, the testing methods that can be used vary dramatically. The testing methodologies used over the years were developed with specific goals in mind, focusing on the need for accurate reliable and replicable measurement of perception. Another issue is the need for bias-free assessment, within a reasonable testing period, given the short attention span of most young listeners. In the late 20th century, behavioral habituation and dishabituation methods were developed, whereby infants are exposed to a stimulus, and once they habituate to it they stop responding. Subsequent responses to novel stimuli represent a demonstrable and measurable indication that the infant has detected a change in sensory stimulation. This technique may ultimately underestimate sensitivity, because conditioned reflexes/responses subside after repeated presentation of novel stimuli. Nonetheless, basic responses have been measured, and are rather useful in underscoring

development of auditory sensitivity in infancy. For example, by 35 weeks' gestation fetuses can apparently discriminate between 250-Hz and 500-Hz tones (Shahidullah and Hepper, 1994). Another approach, that of visual reinforcement audiometry (Moore et al., 1975; Trehub et al., 1981), leads to prolonged responses and may produce better and more reliable threshold estimates with repeated stimulation, because the infant has a motivation to respond. However, the reinforcement method only works once an infant is mature enough to engage in that task (∼5 months of age).

In the 1980s Olsho and colleagues (1987) developed the observer-based psychophysical method (OBPM), to test infants younger than 5 months as well as older infants. The OBPM relies on bias-free ways of determining whether an infant hears or does not hear a given sound; similarly, whether an infant has heard a difference between two sounds. A researcher, i.e., "observer," is "blind" as to the stimuli being presented, and is trained to watch the infant during specific testing intervals. The observer determines whether, based on changes in the infant's behavior, a change in stimulus has occurred. The behaviors that are reasonable to include, in addition to overt head turn and eye gaze, are more subtle behaviors such as eye widening and brow movement. Although not often used with special populations such as persons with cognitive and/or physical disability, this approach has the strength in its objectivity, such that it could theoretically be adapted to any population. With normal-hearing infants, the OBPM can be used to determine thresholds for hearing sounds of various frequencies and other characteristics as early as 2–4 months' age postnatal (Olsho et al., 1987, 1988), and has been applied to testing of toddlers (Grieco-Calub et al., 2008; Grieco-Calub and Litovsky, 2010). However, in newborn infants the method is less reliable (Tharpe and Ashmead, 2001).

Finally, by the time that children reach 3 years of age or older, they can typically engage in tasks involving computerized, interactive platforms. Many of the studies in the past ∼20 years with ages 3+ years have utilized these approaches to study auditory maturation in children. An obvious user-friendly approach is one whereby the child is given options for responding, known as the "forced-choice" methods. For example, if one is measuring the child's ability to perceive a difference between two tones, a two-alternative forced-choice method would provide the child with two types of stimuli and two response options. Experimenters vary the parameters that change and the size of change (large change, easier to hear vs small change, harder to hear). Litovsky (2005) used a four-alternative forced-choice procedure to test speech intelligibility, whereby four pictures of known words appeared on the computer screen and the child's task was to identify the picture that matched the word that was presented from a loudspeaker. By varying the sound intensity at which the word was presented, speech reception thresholds were then estimated, in quiet and in noise. Similar approaches have been used to study the development of sensitivity to important features of sound. These include detection of sounds at the lowest intensity possible, discrimination for stimuli varying in intensity and loudness, frequency and pitch, as well as sensitivity to duration of a stimulus and detection of a gap in a stimulus. Finally, studies focus on effects of masking on threshold for detecting stimuli, on how subjects group together sounds that share characteristics, and on sound localization.

Estimation of sensitivity and threshold in the world of psychoacoustics takes two primary approaches. The first is to design experiments that measure psychometric functions, whereby responses to many stimulus values are collected and percent correct or percent of trials with detection is computed, and from those functions, threshold is defined as the point on the function that meets some agreed-upon criteria that is different from chance (50%, 71%, or higher, depending on the experimental procedure). This approach provides abundant data at all stimulus values (e.g., low to high levels, or small to large differences between stimulus features being discriminated). However, this method can be quite laborious and may not be feasible with young listeners. An alternative, more efficient method involves an adaptive tracking algorithm, whereby the stimulus feature of interest changes following the response of the listener, and is adapted to "track" performance. Thus, when the stimulus values presented result in success, changes are implemented to bring the values to a more challenging value. In contrast, failure to hear a sound or to discriminate between two values of interest results in values that are less challenging. Much work has been done on identifying more efficient and revealing algorithms for achieving adaptive tracking goals. For further readings on this topic please refer to Macmillan and Creelman (2004).

In non-human species behavioral testing is somewhat similar to testing very young human listeners. The animal must be trained to respond to sounds and motivated to do so efficiently and effectively. Typically, studies with non-human animals involve situations in which the animal is deprived of food and/or water outside of the experimental setup, and the animal "earns" its nourishment by producing behavioral responses to stimulation. These studies are typically conducted with fully mature animals, rather than developing ones, because even the mature animal is often very difficult to train. In fact, in most cases, animals are trained for weeks, if not months, before the experimenter is convinced of their ability to perform on the task reliably.

It is fairly accurate to note that mature non-human animals behave much like very young humans. The common characteristic is that they do not share a spoken language with the experimenter through which instructions can be given. Thus behavioral training is an essential component of successful experimentation, and reliable results are ultimately dependent on the integrity with which the listener is trained to understand the task and is motivated to perform. To date, many animals have been studied behaviorally and a comprehensive review of these is outside the scope of this chapter. Finally, as is discussed below, auditory abilities emerge at various times during postnatal maturation. The estimate of the age at which these abilities reach adult level maturity is determined by the behavioral testing done. Table 3.1 summarizes these by indicating the age range during which development reaches adult-like maturity, for aspects of these developmental milestones.

Table 3.1

Summary of auditory development findings

	Infants <6 months	Infants <12 months	Toddlers <3 years	Preschool age	Early School age	Middle School age	Teenage years
High Frequency Detection	←→						
Low Frequency Detection					←→		
Intensity Discrimination					←→		
Frequency Discrimination		←→					
Loudness Perception		←- - - - - - - - - - - - - - - -→					
Simultaneous Masking					←- - - - - -→		
Auditory Scene : Streaming						←→	
Co-modulation Masking Release				←- - - - - - - - - - - -→			
Informational Masking				←- - - - - - - - - - - - - - - - - - -→			
Backward Masking					←→		
Spatial Release from Masking			←→				
Contralateral Un-Masking					←→		
Minimum Audible Angle (MAA)					←→		
MAA with echo (precedence)							←→

Continuous arrows, age range during which perceptual ability is known to be mature; dashed arrows, age range during which perceptual ability is somewhat mature but additional research may be required.

CODING OF AUDITORY FEATURES

While behavioral responses to sound can be seen in infants as early as 28 weeks' gestation (Birnholz and Benacerraf, 1983), the ability of humans to perform on behavioral auditory tasks matures into the late teenage years (Maxon and Hochberg, 1982; Fischer and Hartnegg, 2004). As discussed below in greater detail, the auditory system undergoes protracted maturation, which is one of the mitigating factors in auditory development. However, non-sensory factors, including attention, cognition, and memory, are known to play a crucial role in maturation of auditory perception, and in the individual variability that is seen within age.

Detection of sound

Much of the early work on the stimulus level required to detect stimuli was done on auditory sensitivity in quiet, that is, in absence of noise or other masking stimuli such as speech. For example, it was well documented that thresholds for detecting stimuli in quiet improve between infancy and early school age (Olsho et al., 1988; Trehub et al., 1988). Infants can have thresholds up to 25 dB worse than adults; rapid improvement occurs before the age of 6 months (Tharpe and Ashmead, 2001), with a difference narrowing down to a 10–15 dB gap by the time that children reach 5 years of age. Differences in detecting sounds seem to also depend on the frequency, whereby sensitivity is near adult-like by age 6 months for high-frequency sounds of around 4000 Hz, but not for low-frequency sounds around 500 Hz (Olsho et al., 1988; Trehub et al., 1988). The physiologic basis for this difference has not been identified, and does not appear to result from maturation in conductive properties of the peripheral auditory system or in the sensory development that takes place during infancy and childhood. More likely is the fact that the tasks used to measure perception do not separate out effects of cognitive ability, motivation, memory and variability in neural representation of the stimuli, otherwise known as internal noise.

Frequency and intensity discrimination

Frequency discrimination, which refers to the ability to perceive a change in the frequency of tonal stimuli, undergoes considerable maturation between 3 and 6 months of age. Testing for this ability is done by presenting a train of tones, and switching to a tone of different frequency midstream, then noting whether the infants' behavior changed. While discrimination ability is poorer for 4000-Hz tones than 500-Hz tones at 3 months, at 6 months the pattern is reversed and infants are better at discriminating changes imposed on 4000-Hz tones than 500-Hz tones (Olsho et al., 1987). Remarkably, adult-like performance of discriminating ~1% change in the frequency is reached between 6 and 12 months of age for 4000-Hz tones; infants can discriminate changes of 2% in the frequency. At 4000 Hz that would mean that infants can perceive a change of 8 Hz (4000 Hz–4008 Hz). Interestingly, measures of frequency discrimination in children suggest that thresholds are higher than those seen in infants, but the tasks used with children are notably different, and likely place a heavier memory and attentional load on the child than the infant task (Moore et al., 2008). Thus, an important conclusion is that basic auditory abilities may appear to be highly developed in infancy, if tested using simple tasks that select for measuring sensation rather than attention or memory.

Intensity discrimination refers to the ability to detect a change in the level (in dB sound pressure level) at which a sound source is presented, and this ability is typically measured separately for narrow ranges of frequencies (tones or narrowband stimuli). Intensity processing is thought to reflect a change in the rate of firing of auditory neurons and possible "recruitment" of auditory nerve fibers that are anatomically neighboring those at which the stimulus is presented. Similar to frequency sensitivity measures, intensity discrimination can be tested by measuring the ability of a listener to detect a change from a background stimulus. Adult listeners can hear differences in intensity as small as 1–2 dB (Sinnott and Aslin, 1985), and, in contrast, infants between the ages of 5 and 7 months need approximately 6 dB difference. As far as stimuli are concerned, infants are particularly worse than adults for low-frequency stimuli (~400 Hz) than high-frequency stimuli (4000 Hz), and for stimuli that have a standard (sound to which changes are compared) that is low in intensity to begin with. There are clear methodologic issues here related to how exactly sensitivity is measured, and whether the task is difficult or much more subtle for infants. This issue is not exclusive to intensity discrimination, but may be appear to be more relevant because there are not many demonstrations of mature ability for intensity discrimination in infancy. In fact, it may be the case that maturation of perception for changes in intensity continues into childhood (Buss and Hall, 2009). It is possible that cognitive abilities, including memory, play an important role in the tasks used to date; however, tasks that lead to equal performance between very young listeners and adults have not been identified, suggesting that intensity discrimination is more immature than frequency discrimination early in life.

Loudness

Another aspect of auditory perception that relates to sound level or intensity is that of perceived loudness of a stimulus. Whereas discrimination, whereby we measure whether a listener was correct when s/he reported whether a stimulus changed or did not change, is objective, loudness perception is subjective. Any subjective perception is difficult to measure in listeners, including not only infants and children, but adults as well. Loudness is an attribute for a sound that places perception on a scale ranging from inaudible/quiet to loud/uncomfortable, in response to change in sound pressure level (intensity). Because there is no correct answer, there is some challenge in knowing when and how to reinforce a child and how to train the child to respond. Nonetheless, it appears that, while some children have difficulty learning the task, others can perform similarly to adults (Serpanos and Gravel, 2000). Because loudness growth is abnormal in people with hearing loss, such that loudness grows rapidly over a small range of intensities, emphasis on understanding the importance of loudness perception maturation may come from the audiologic literature, with hearing-impaired children. Evaluations of hearing-aid fittings and perceived loudness of speech signals after amplification become clinically crucial, so that the speech signal is heard, understood, and comfortably presented (e.g., Scollie et al., 2010; Ching and Dillon, 2013). Future work on basic psychophysics may be important in order to capture the perception of young infants and children with typical hearing, and benchmark their abilities, so that expectations are appropriate for children who are fitted with hearing aids.

MASKING AND AUDITORY SEGREGATION

Background on grouping and segregation

Natural auditory environments such as classrooms, home environments, entertainment areas, and other public space are typically complex in terms of sensory stimulation. In the auditory domain the listener is faced with a cacophony of information that varies in content, location, identity of the source, frequency, intensity, and other; these features are also typically dynamic rather than static, thus they are likely to change over time. It is the infant's or child's job to learn how to cope with these environments, to organize the onslaught of stimulation into perceptible nuggets of information, so that the child can learn about his or her environment, attain language capacity, and master numerous abilities that rely on auditory-based communication.

At the initial stages of auditory perception, as described earlier in this chapter, the auditory system masters the ability to detect and discriminate between sounds. A further stage of analysis requires that decisions be made regarding which sounds "belong together" and ought to be "grouped," and which sounds "do not belong together" and ought to therefore be "segregated." This issue is not so much about categorization, but rather this is ultimately about identification of sounds and extraction of meaning from sound sources. Categorization might be relevant when determining which sounds are emitted by humans vs non-humans, or deciding whether sounds arise from musical instruments vs natural environmental sounds (a stream or the sound of the wind rustling leaves on a tree). In contrast, identification and information extraction require that the listeners be able to segregate a sound source from other sounds in the environment. For example, a child who is attempting to listen to her mother's voice in a crowded room must separate that voice of interest from other people's voices and other sounds in the environment. In some sense, the "other sounds" can be thought of as competitors or interferers; in the classic psychoacoustic literature they are also referred to as "maskers." The basic process through which the neural mechanisms involved in perception are able to parse out various sound sources and assign meaning to appropriate sound sources has been termed auditory scene analysis (ASA: Bregman, 1990). A more specific example of ASA that occurs for speech sounds is the cocktail-party effect (Cherry, 1953). Bregman's notion was consistent with two stages – one that is more "primitive" and does not require much learning or experience, and a second stage that is "schema-based" and is organized through each listener's experiences and learned abilities during a lifetime of listening. Most of the work done in the field in general has been with adult subjects. Developmental work in this area has primarily attempted to understand the "primitive" case for ASA, because the lack of dependence on experience means that we can begin to understand mechanisms involved in the organization of incoming stimuli that are existent in young listeners.

To the extent that these mechanisms exist at a young age they can potentially serve as the scaffolding upon which more complex schema-based organization of information can occur with experience. This issue of course lends itself to fascinating questions about ways in which deprivation of stimulation during development results in disruption of the primitive organizational structures and/or in ways in which schemas can in fact develop properly. Very little is known about ASA and effects of deprivation during development on the ability

of listeners to organize incoming information once the auditory system has been reactivated.

One reason that little is known about this issue is that there are no examples of cases in which human listeners have undergone auditory deprivation early in life and have had natural hearing restored henceforth. The closest that we can get to a "natural" experiment of this sort is to study effects of auditory deprivation in people who are born deaf and who are able to receive auditory stimulation later in life through artificial means. Today's technologic advances offer the cochlear implant (CI), which provides hearing through electric stimulation of the auditory nerve (Wilson and Dorman, 2008). An important caveat is that the CI does not restore normal hearing to recipients; rather, the auditory signal is compromised by spectral degradation, poor specificity of stimulation along the tonotopic (frequency-dependent) mapping of the auditory system, and neural death following deprivation, to name a few limitations (Shepherd and Hardie, 2001; see Litovsky et al., 2012). Thus, interpretation of data on how deprivation might affect ASA are unavoidably compromised by the inability to know whether poor performance following auditory stimulation with a CI is attributable to the CI-related issues or to factors related to the deprivation *per se*.

Energetic masking

The concept of energetic masking was introduced when another form of masking, known as informational masking, became prominent in the literature (see below). Energetic masking is thought to be the interference produced by noise or other stimuli that can vary in characteristic, but that share features with the target sound, and can be accounted for by considering the peripheral auditory system. A simple example of energetic masking is when a noise burst occurs simultaneously with a tone, and the two-frequency bandwidth of the noise is centered around the frequency of the target zone. In energetic masking, one sound interferes with our ability to detect or otherwise hear another sound because the two sounds excite similar auditory neurons in the periphery, thereby limiting the extent to which information about the target can be perceived in the presence of the masker (Fletcher, 1940). Several studies have examined the ability of infants and children to detect a signal in background noise, either when the sounds are presented simultaneously, or when they are temporally offset. For simultaneous masking detection, thresholds of infants aged 6–24 months are higher than those of adults by l5–25 dB (e.g., Bull et al., 1981; Bargones et al., 1995), suggesting that their ability to extract a simple signal in background noise is worse than that of adults. Similar to

the thresholds discussed above for detecting signals in quiet, there appears to be a more rapid maturation for hearing signals in noise, with high-frequency thresholds reaching adult-like performance sooner than low-frequency thresholds (e.g., He et al., 2010).

Auditory streaming

The field of perception has broadly defined relevant terms that address the fascinating problem of ASA. For example, auditory streaming refers to the notion that when listeners are presented with sounds that share some dimensions and vary along other dimensions, they perceive them either as one coherent sound, or as two distinct sounds. A classic example of auditory streaming is when two sequences of tones that vary in frequency are presented; the more similar the frequencies, the more likely it is that listeners will perceive a single, coherent auditory image. Auditory stream formation appears to depend on acoustic parameters, including frequency, spatial locations, sex of the talker, spectrum, and common temporal onsets/offsets (Bregman, 1990; Yost, 1997). Using a method whereby infants are habituated to a signal, then dishabituated to a novel stimulus, it has been shown that by the age of ∼4 months infants can use similar acoustic cues to those used by adults to perceive two stimuli as either belonging to a single stream or to two streams (McAdams and Bertoncini, 1997). It is important however to note that the sensitivity to these cues was not measured with these studies, thus it remains to be understood whether infants need the same amounts of cue similarity or difference in order for the effect to be operative.

In contrast to studies with infants, by the time that children are 4–5 years of age, methods for testing auditory perception can be fairly similar to methods used in adults. When children and adults were tested on their ability to hear one or two tone sequences, developmental changes have been observed between children who are 8 years old or younger and children ages 9–11 years. Older children, like adults, required small differences between the frequencies of the two tone sequences in order to perceptually segregate them into two "streams" whereas younger children required larger frequency differences in order to hear the tone sequences as segregated (Sussman et al., 2007). These findings suggest that ASA, as measured with auditory streaming, develops well into school-age years, as will be discussed below for a number of other auditory abilities.

Co-modulation masking release

Auditory grouping can also be studied using a completely different paradigm, known as co-modulation masking

release (CMR), which was originally described by Hall et al. (1984). This fascinating perceptual phenomenon can be elicited under conditions when a target signal is masked by another sound that has intensity fluctuations known as amplitude modulations. When a masker at a frequency that is different from that of the target signal is added, and has the same temporally varying modulations in intensity, i.e., when the target signal and added masker are "co-modulated" in time, the two sounds are perceptually grouped and perceived as a single auditory image. This grouping leads to benefits in signal detection, i.e., reduction in listeners' thresholds for hearing the target signal. Studies on CMR in children suggest that, although children experience more masking in the presence of noise compared with adults, when the added co-modulated noise is introduced, they show similar amounts of benefit as adults. Thus, CMR is adult-like under certain conditions, suggesting that mechanisms involved in grouping under noisy conditions, and the ability to benefit from the temporally co-modulated noise, are developed by 4–5 years of age (Veloso et al., 1990; Grose et al, 1993). Other studies have shown that CMR can also undergo more protracted maturation, for example, when the target signal and noise are both modulated, but with a temporal asynchrony between them. In that scenario, CMR in children was eliminated, whereas in adults the effect was reduced relative to the condition with temporal synchrony (Hall et al., 1997).

Informational masking

An important concept in the context of how it is that we extract information in complex auditory scenes is that of informational masking. In contrast to energetic masking (see above), informational masking refers to masking that cannot be accounted for by peripheral mechanisms, for example when a tone is difficult to hear despite the fact that the masking noise is presented at a frequency region that is remote and that stimulates a different population of auditory neurons. A simple example of informational masking is when a series of "off-frequency" tones change in frequency through time in a manner that is unpredictable each time they are presented. Because the listener does not know which aspects of the spectral, temporal, or other features of the masker to ignore, uncertainty occurs, which results in increased threshold for the target tone. In other words, informational masking might simply indicate that the listener is confused about which features of the stimulus belong to the target and which to the masker (Durlach et al., 2003).

Studies on informational masking began to get attention 30–40 years ago (Watson et al., 1976; Neff and Green, 1987), when several paradigms were implemented using tone bursts to study the ability of adults with normal hearing to hear tones in the presence of multitonal background maskers, or "interferers," whose content was varied in an unpredictable manner. By randomizing the frequencies of the tones from presentation to presentation, researchers introduced uncertainty about what listeners should ignore. This was quite unlike, then traditional, studies with energetic masking, where the maskers are consistent across presentations, and listeners know what sound to ignore and which sound to attend to. Informational masking was defined then as the amount of masking observed, above and beyond that which was observed with energetic masking conditions. This paradigm resulted in demonstrating vast individual differences (Oh and Lutfi, 1998). In addition, there is an interesting non-monotonic relationship between the amount of masking and the number of tones, such that, with a small number of tones, as more are added masking increases, suggesting that the confusability or uncertainty may be related to attention or memory load. However, at some stage the increase in number of tones results in the tone complex being perceived as a noise with little information, hence the amount of masking begins to drop (e.g., Neff and Green, 1987; Oh and Lutfi, 1998).

This paradigm has been successfully applied to studies with infants and children and has led to important findings regarding auditory development. In general, studies suggest that infants and children are more vulnerable to effects of informational masking, that is, the amount of masking they experience is larger than that seen in adults (e.g., Oh et al., 2001; Leibold and Werner, 2006). Lutfi et al. (2003) investigated informational masking using tone bursts in 38 children aged 4–16 years, and focused on identifying potential sources of variability to account for different strategies that children might use to perform on this task. The authors concluded that, under easy-listening conditions, children generally utilize what is in the world of psychoacoustics known as an "optimal" decision strategy to solve this problem. However, in the informational-masking paradigm, differences between children existed at all ages, as is commonly found in other studies on auditory development. Moreover, these were thought to reflect the extent to which uncertainty in the masker affects the ability of the child to "focus in" on the signal. However, the authors did not discuss the possibility of utilizing this tool for ultimately measuring auditory-based, but not speech- or language-based, attention and memory in children.

Leibold and colleagues have also examined effects of uncertainty when masker spectra do not vary, and found developmental differences, suggesting that infants and children are perhaps generally poorer at segregating

the target from masker, even when there is uncertainty regarding "what to ignore" (Leibold and Werner, 2006; Leibold and Neff, 2007). This clearly suggests that the conclusion regarding developmental effects depends on exactly how the experiments are set up and what the stimuli and tasks consist of. Nonetheless, the fact that developmental findings are clearly found suggests that there are reasons to be concerned about the ability of children to listen in complex, realistic acoustic environments, where sound sources vary in many features, including location, number, time, spectrum, and amplitude. More emphasis on effects of location and reflections is placed on sections below, where spatial hearing is discussed. In general, when it comes to speech sounds and the effects of informational vs energetic masking, again the issue revolves around the extent to which the masker is similar to the target, and thus the extent to which the two can be confused with one another. For example, the voice of one person can be easily confused with that of another if the two are similar, and if the content overlaps. In contrast, different voices, such as those of a young child (high pitch) uttering a nursery rhyme vs a male with low fundamental frequency talking about the stock market, are less likely to be confused.

Furthermore, there is the issue of whether content can be extracted from the maskers vs whether it sounds like noise. Newman (2005) found that infants can recognize known words in the presence of multitalker babble, which is more similar to noise and has less discernable content than speech, and in contrast they failed to do so in the presence of a single "masking" talker, perhaps because its content was more distracting and thus confusing. Finally, there is the added issue of whether the sounds are familiar to the infants. Barker and Newman (2004) found that when the target voice was their own mother (thus familiar), infants were better able to understand speech in the presence of background talkers compared with a situation in which the voice was not familiar. Performance is also measured by obtaining the signal-to-noise ratio (level in decibels of the target re: the masker) that is needed for the target to be understood. Compared with adults, infants generally require a higher signal-to-noise ratio than adults in order to recognize their name in the presence of a noise masker (Newman and Jusczyk, 1996). These level differences are generally higher for children than for adults (Elliott et al., 1979; Nittrouer and Boothroyd, 1990; Hall et al., 2002; Litovsky, 2005).

These studies are also related to the issues described above in relation to the cocktail-party effect, and raise the issue of whether children are simply poorer at selectively attending to the target stimulus in the presence of the masker. Some of the more interesting studies in this area are done with event-related potentials, which measure physiologic responses in the central auditory pathways. For example, when adults and children were each asked to attend to one of two passages read to them, the brain activity as measured with the N1 wave was not adult-like at age 6–8 years (Coch et al., 2005). While somewhat remote from neural mechanisms involved in the task, one might argue that lacking characteristics that are adult-like in brain waves may reflect the immaturity that is observed in behavioral studies. It is not surprising that neural maturity for central processing of complex information continues to mature beyond ages 6–8 years. Cortical maturation is known to extend into the late teenage years, and effects of experience are certainly bound to have significant effects on these maturational changes.

Backward masking and auditory maturation

As discussed below, one way to study masking is to present signal and noise at the same time and to vary few parameters. Thus, the interference from the noise can be easily accounted for by peripheral mechanisms. In contrast, backward masking is studied by presenting the signal before the noise; hence, a more central mechanism that weights information arriving from the signal and noise must be considered. In fact, it appears that, unlike simultaneous masking, which is adult-like by age 6–10 years, backward masking results in significantly worse performance in children than adults. It has been suggested that temporal resolution is involved in backward masking and is an ability that continues to refine during childhood (Hartley et al., 2000). The physiologic mechanisms underlying these developments are not well understood, but it has been suggested that peripheral frequency resolution cannot explain these findings (Spetner and Olsho, 1990).

Backward-masking paradigms have been ubiquitous in studies with populations of children who demonstrate language delays and reading disorders. It appears that maturational delays occur when the brain processes temporal information. For example, there appear to be delays in maturation on tasks that measure backward masking and on other masking paradigms in which there are delays between the signal and masker. It has been argued that language deficits may result from limitations in auditory processing (Wright et al., 2000), which may have implications for diagnosis and/or remediation in children with language impairment or other disorders with a related underlying mechanism. Wright et al. (1997) argued that brains of individuals with language delays and reading disorders develop at a slower rate than brains of their unaffected peers. They honed in on the age of ≈ 10 years, when brain changes associated with

onset of puberty could affect perceptual development. Numerous other studies have identified interactions between delayed brain maturation and perceptual and cognitive abilities and suggest that puberty may be an important factor contributing to learning problems (Doupe and Kuhl, 1999; Bourgeois et al., 2000).

Finally, the field of rehabilitation, training, and remediation for those affected with learning disorders is robust and the treatment options are beyond the scope of this chapter. One interesting point, however, is whether effects of treatment or training can be captured in the neural representation of cues that are important for perception, including speech cues. There appears to be some evidence for short-term training effects on improving neural representations of auditory cues that are known to be important for understanding speech in noise (i.e., masked speech). This evidence comes from measurements of biologic responses at the level of the brainstem (Song et al., 2012).

SPATIAL AND BINAURAL HEARING

Some of the most robust effects from listening in noise, and the ability to overcome challenges that arise in noise, are observed when listeners have the capacity to utilize input arriving at both ears. Thus the cocktail-party effect has been widely studied in relation to what are known as binaural and bilateral benefits. Related to these benefits is the ability to locate where sounds are coming from, which also depends greatly on acoustic cues known as binaural cues. Finally, as discussed below, the attainment of a normally developed sound localization system depends on the integrity of the auditory system's ability to extract spatial cues and present them to the auditory system with fidelity.

Binaural cues

In order to consider the development of spatial and binaural hearing, one must first consider the acoustic cues that arise when sound sources occur from various locations in the environment. Sources arriving from the side will reach the two ears with differences in time of arrival and intensity. For example, a sound at 90° to the right will produce a difference of ~0.7 ms in interaural timing, otherwise known as the ITD. This cue is robust at low frequencies, typically below 1500 Hz. At high frequencies the head creates an acoustic "shadow," so that the near ear receives a greater intensity than the far ear; sounds are thus received at the two ears with interaural level differences, otherwise known as ILDs, and these can be as large as 20 dB. For sounds that are amplitude-modulated, such as speech, ITD cues are also available from differences in the timing of the envelopes (slowly varying amplitude) of the stimuli. Figure 3.1 provides a schematic of the directionally dependent cues
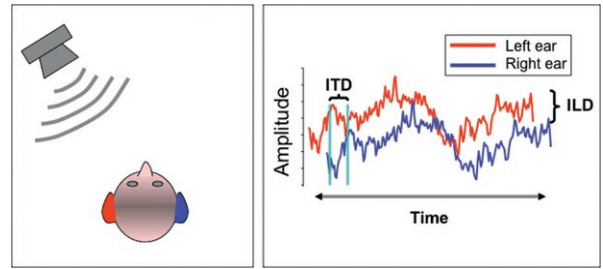

Fig. 3.1. Schematic of the directionally dependent cues that would be potentially available to listeners in the horizontal plane for a broadband sound. The left panel shows the sound emitted from a loudspeaker and arriving at the left ear first and with greater intensity. The right panel shows the measurements at the ear canal of the two ears; both interaural time difference (ITD) and interaural level difference (ILD) cues are present. (From "Bilateral Cochlear Implants in children" by Ruth Y. Litovsky and Jane Madell in *Clinical Management of children with cochlear implants* (P. 61) by Laurie S. Eisenberg. Copyright © 2009 Plural Publishing Inc. All rights reserved. Used with permission).

that would be potentially available to listeners in the horizontal plane for a brief signal such a click.

Measuring localization

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 numeric 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 (Middlebrooks and Green, 1991).

With children, responses regarding perceived source location are challenging to obtain, and few studies on this topic exist. However, similarly to studies on other aspects of auditory signals described above, localization prediction has been measured in infants and children using the discrimination paradigm. A common measure 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). It is important to note that 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; Grieco-Calub and Litovsky, 2010).

Developmental findings on localization

Newborn infants orient towards the direction of auditory stimuli within hours after birth. This head-orienting response is unconditioned, or reflexive, and seems to enable the infant 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 3.2 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).

Children between the ages of 4 and 10 years can localize sounds with average error rates ranging from <5° to >30°. Grieco-Calub and Litovsky (2010) found that

RMS errors ranged from 9° to 29° (average of 18.3° ± 6.9° 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° ± 10.72° SD), which overlapped with that observed with adults. Van Deun et al. (2009) found that RMS errors averaged 10°, 6°, and 4° for children aged 4, 5, and 6 years, respectively. The question is whether sound localization using this type of source location identification paradigm can also be measured in very young children. Recent work by Litovsky and colleagues suggests that 2-year-old children can respond to sounds arriving from one of nine locations (Ehlers et al., 2013). Interestingly, most children identify locations correctly on >95% of trials (RMS errors <10°), while a small number of children identify incorrectly more often (RMS errors nearly 30°). These findings are summarized in Figure 3.3.

Across these studies it is also clear that children vary in their ability to perform tasks of absolute localization. While 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 to non-sensory factors, as discussed above for other findings in perceptual development, needs to be further explored. 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.

There is evidence from the animal literature to suggest that the auditory cortex plays a key role in determining the ability of an animal to localize, and to learn to localize or to relearn novel maps of space. Studies in which the cortex was ablated, conducted over 40 years ago, led to fairly impaired performance, but the damage to the cortex was difficult to control (Whitfield et al., 1972). More recently the ferret has been used as an
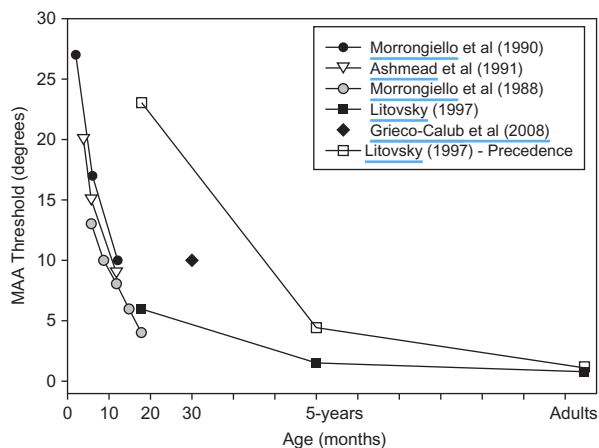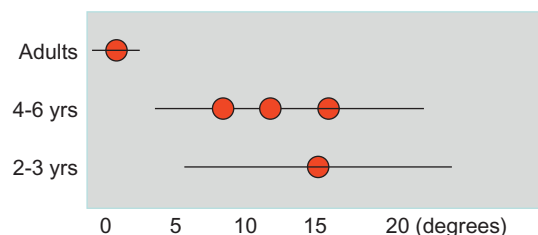

Fig. 3.2. Developmental trend in right–left discrimination ability, as measured with head-orienting responses. Data represent minimum audible angle (MAA) thresholds from several studies in which the sound intensity was held constant from trial to trial. From these studies, one can conclude that the largest decrease in MAA occurs between 2 months and 2 years of age, with continued improvement through 5 years of age, when children's thresholds are similar to those of adults. All studies by one used stimulus levels that were fixed. The diamond at 30 months represents data collected with stimulus levels roving an 8 dB range. (From Litovsky, 2011, with permission from Springer Science and Business Media).

Fig. 3.3. Root-mean-square (RMS) error values collected in various studies are shown here, with age on the y axis and mean (±range) of RMS values measured. This is intentionally vague to represent the range of abilities tested, because the measures depend greatly on the exact stimulus used, loudspeaker configuration and task.

animal model of the effects of unilateral occlusion on spatial hearing. Most notable are two findings. First is that lesions of the auditory cortex (the primary auditory cortex (A1) in particular) lead to a reduction in the ability of animals to relearn spatial hearing maps (Nodal et al., 2010, 2012). Second, their behavior improves most dramatically with training and feedback (Kacelnik et al., 2006), suggesting again that non-sensory factors are likely to be involved in the emergence and preservation of spatial hearing maps.

Binaural unmasking

In this section, the issues related to complex auditory environments and ASA are reintroduced. However, the focus here becomes the availability of spatial and binaural cues for segregating target speech from competing speech and/or noise. Normal-hearing listeners rely on 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). Average data from several studies are shown in Figure 3.4; these values can be quite large (10–12 dB) or relatively small SRM (3–5 dB), depending on the stimulus of choice, type of masker, and task. 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., informational masking: see above). The magnitude of SRM with speech-on-speech masking can be quite large relative to noise maskers (Durlach et al.,

Fig. 3.4. Range of spatial release from masking (SRM) data collected from numerous studies is shown. Age on the *y*-axis and mean (±range) of SRM values measured with numerous studies is shown. This is intentionally vague to represent the range of findings, which depend greatly on the stimulus used, choice of words, and sentences, and whether informational or energetic masking occurred.

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 conditions in which the monaural "head shadow" occurs.

Litovsky (2005) first demonstrated SRM in children aged 4–7. Target stimuli were presented from the front at 0°. Maskers were presented from locations that were either co-located with the target, or spatially separated from the target. For children, a four-alternative forced-choice task is implemented in order to enable children to indicate what they hear rapidly and reliably. In the co-located condition children experience masking that is greater than that seen in the separated condition. The difference between the two conditions, or SRM, represents the ability of the listener to segregate sources based on location. SRM reported by Litovsky (2005) was 5–7 dB improvement in speech reception threshold. This test is an excellent tool for delineating children who are unable to utilize spatial cues for source segregation, as demonstrated by Misurelli and Litovsky (2012) with CI users. Garadat and Litovsky (2007) were able to show the effects in 3-year-old children, and more recently, Hess (2013) found SRM that was fairly mature in 2-year-old children. One key factor is to select target words that are within the receptive language and vocabulary of the children being tested.

Finally, as was discussed above in regard to informational masking, SRM is largest when target masker stimuli are more easily confusable with one another, such as if they comprise the same sex vs different sex talkers (Johnstone, 2006; Misurelli et al., 2013). Source segregation has also been studied over headphones, with a focus on non-speech stimuli. Thus, a related measure of the binaural masking level difference (BMLD) provides vital information regarding separation of a target signal such as a tone or narrow band from 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 co-located intracranially in the N_0S_0 condition, while they are perceived as spatially separated in the N_0S_π.

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). Nozza et al. (1988) reported BMLDs for speech stimuli of 5, 8, and 10 dB for infants, 4-year-olds and adults, respectively, and suggested that there are developmental effects due to changes over time in absolute sensitivity to signals presented in masked conditions. Other studies, with children between
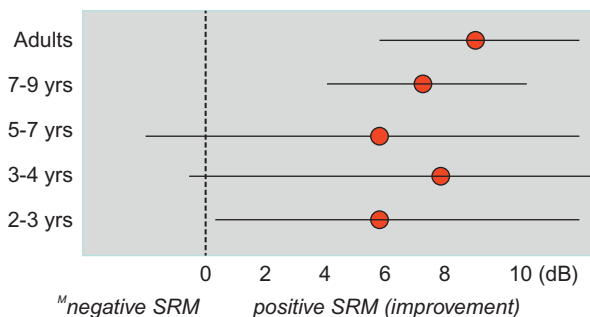
the ages of 4 and 12 years, have shown that BMLDs are similar in children and adults for wideband noise bursts. However, there are age effects when narrowband noise is used, such that children ~4 years old are generally slightly worse than older children or adults (~7–12 dB depending on the noise bandwidth; Grose et al., 1993; Van Deun et al., 2009). It is possible that BMLDs are related to the ability of the listener to integrate information through time, and that this temporal window develops with age. Thus, adults might 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. In contrast, 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 larger time windows (Hall et al., 2007).

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 (Fig. 3.5A). In reverberant rooms, although listeners are aware of the presence of reflections, localization cues carried by reflections are de-emphasized 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, henceforth referred to as the precedence effect (PE) (Blauert, 1997; Litovsky et al., 1999). While 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.

As is illustrated in Figure 3.5B, stimuli in PE studies typically contain a simulated, single source and a simulated single "echo" or reflection, known as the lagging sound. 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
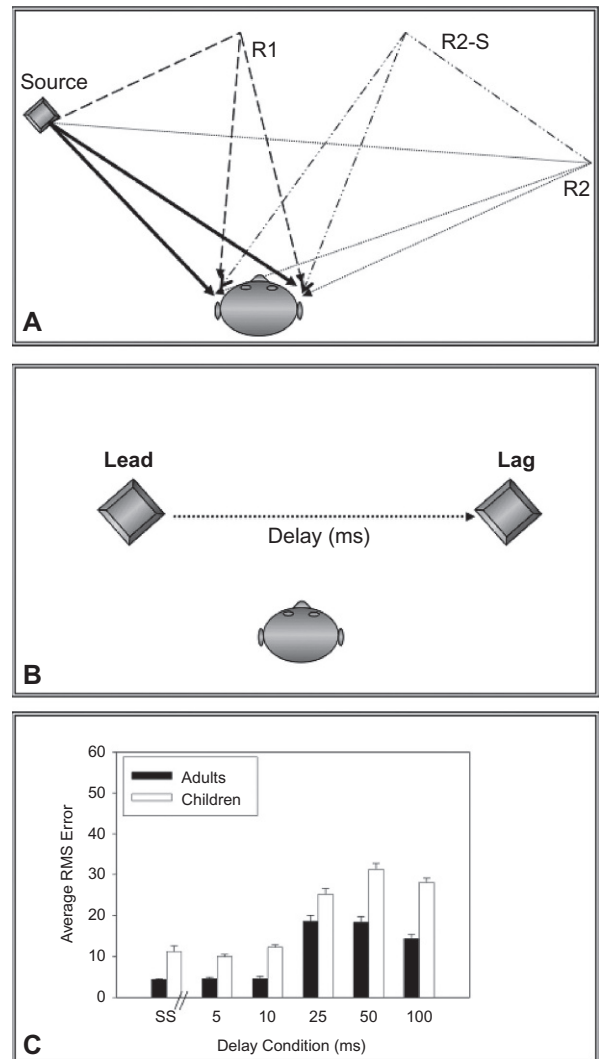

Fig. 3.5. (**A**) Rendition of the source and reflections that a listener might encounter in a room. Sounds reach the ears through a direct path and through reflections from the walls. (**B**) Stimuli used in precedence effect experiments, which typically contain a simulated source (lead) and a single "echo" or reflection (lag), thus a much more simplistic version than the ones shown in (**A**). (**C**) RMS, root mean square. Data from Litovsky and Godar (2010). (Reprinted with permission from Litovsky RY and Godar SP. Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. J Acoust Soc Amer 128 (4) 1979–1991. Copyright (2010), Acoustic Society of America.)

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 similar fashion to reduced fusion.

In newborns, the use of the reflexive unconditioned head turns has been used, with the lead stimulus to one side and the lag towards the other side; a listener

who has functional PE would hear the image on the side containing the lead – "localization dominance." As reported by Clifton et al. (1981), infants do not orient their heads towards the location of the leading source, suggesting that lagging sounds or "unheard echoes" can exert a strong influence on directional hearing in young listeners. At age 4–5 months localization dominance seems to have developed (Muir et al., 1989); however, the strength of echo suppression continues to mature into middle to late childhood. 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 Figure 3.2, where the data from Litovsky (1997) in the PE condition are plotted along those from single-source measurements. The fact that MAA thresholds in PE conditions exceed those of single-source conditions, for the children but not adults, is an indication of the maturation of neural circuits involved in mediating perceptual weighting of sources and their reflections. 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 absence of the echo, MAA acuity is mature by age 5 years. Spatial hearing acuity in children is therefore vulnerable to degradation by virtue of the fact that the auditory system is unable to dismiss directional cues carried by the delayed stimulus.

More recently, Litovsky and Godar (2010) studied localization dominance (the extent to which the leading source dominates sound localization) in children aged 4–5 years and adults. Figure 3.5C shows localization data with RMS errors for localizing the lead, or lag. RMS errors for lead localization were fairly similar to errors observed with single-source stimuli, when the lead lag delays were 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 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 perceptual weighting of spatial cues arriving from sources and their reflections. What factors could account for immature PE in children and infants? Sensitivity to simple (i.e., single-source) 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 (Litovsky, 1998). Finally, it is important to keep in mind the fact that simple stimulus scenarios described here and used in studies of the PE do not represent the "real world" because here only a single reflection is simulated. In addition, in PE studies the reflection has a similar intensity to that of the original sound source. However, the disruption to localization by a single simulated echo is indicative of the potential difficulties that children might experience in realistic, reverberant listening situations.

RELATIONSHIP BETWEEN AGE OF DEVELOPMENT AND DESIRABLE AGE OF INTERVENTION IN DEAF CHILDREN

The primary purpose of providing hearing aids and/or CIs to young children is to habilitate language acquisition, to provide the signal processing necessary for hearing speech in noise, and to provide each individual with the skills needed to function in a mainstream environment. The poignant question thus becomes: "what is the age at which auditory stimulation can minimize loss of auditory function and maximize the ability of the individual to access sound?" In other words, the concern is whether there exists a "critical period" for providing sound, beyond which the child will not be able to maximize his or her potential. On the other hand, even with significant hearing loss and lack of recovery of auditory function, children can function very well in an auditory environment. The answer is complex, because the outcomes in each individual can be quite different, and depend on factors that are not always well understood. The acclimatization strategies vary, and include multimodal use of speech cues, such as lip reading; top-down processing that enables hearing-impaired people to fill in the gaps where speech information is blurred; and numerous other possible strategies.

It is well known that the auditory system can undergo significant changes due to lack of sensory input, resulting in loss of function. This degeneration due to lack of stimulation occurs both peripherally and centrally (Shepherd and McCreery, 2006). At fairly peripheral levels in the auditory system there is known to be degradation in size and function of neural ganglion cells following a prolonged period of auditory deprivation (Leake et al., 1999). Profound deafness in the early developmental period seems to result in loss of normal

tonotopic organization of primary auditory cortex, although there is some reversal following reactivation of afferent input (e.g., Kral et al., 2009). In addition to neural degeneration there is reassignment following deafness (Kawano et al., 1998). There is some evidence to suggest that outcomes in deaf individuals are maximized if stimulation is provided at an early stage in language acquisition, i.e., by approximately 18 months of age (e.g., Niparko et al., 2010).

Regarding spatial hearing mechanisms, there is important application in the area of single-sided deafness, since unilateral hearing loss is prevalent amongst children. In bilaterally deaf children, the issues raised above regarding bilateral implantation become relevant as clinicians determine whether children should receive a second CI at a young age or hold off until the child is older. From research in animal models (e.g., Kral et al., 2013) and humans (e.g., Gordon et al., 2013), there is evidence that early auditory deprivation in one ear, and prolonged stimulation in the other ear, results in aural preferences that are perhaps irreversible. Auditory plasticity though has two important ingredients. First, the level within the auditory system at which studies are conducted is crucial. Peripheral mechanisms are in general more susceptible to deprivation, and less likely to recover function. In contrast, central auditory mechanisms are susceptible to deprivation, likely to show dysfunction due to hearing loss, but also more amenable to plasticity and some recovery of function. Thus, while auditory deprivation can lead to damage in the structures responsible for early stimulus analysis, cortical development follows more iterative interactions and depends on stimulus-driven learning. This means that there is tremendous potential for recovery of some function with training and habilitation. However, it is important to bear in mind the fact that early access to sound does not guarantee good auditory function, because the brain mechanisms must be stimulated and utilized in the growth and learning process (Flexer, 2011).

Finally, windows of developmental plasticity, and ages beyond which the nervous system loses robust responses, are much easier to study with control in non-human species. In humans, studies are conducted retrospectively, by examining children after they have been treated clinically. Thus, combined efforts across the non-human and human approaches can be instrumental in determining best outcomes for children in the long run.

CONCLUSIONS

In this chapter we considered auditory development in the context of the need of humans to achieve auditory perception with fidelity, and interact and learn in a complex world, surrounded by multiple sounds that arrive from different locations. The auditory system reconstructs the perceptual space, groups together some objects, and segregates sounds from one another. In order to engage in an analysis of the auditory world, features of sounds are extracted, and the development of these abilities is staggered throughout development. Auditory development is a broadly defined term, and many approaches have been taken to study this fascinating topic. Here we focused on behavioral studies, which depend on extracting information from listeners who are performing psychoacoustic tasks. These approaches provide a window into the extent to which underlying structure and function support perception. To a large extent, the discussion here focuses on the fact that both sensory and non-sensory, cognitive or listening strategy-driven factors account for developmental changes, and differences between children and adults. Thus interpreting developmental "delays" must be done with caution, and must allow for separation of sensory and non-sensory factors.

ACKNOWLEDGMENT

References

Abdala C, Keefe DH (2011). Morphological and functional ear development. In: LA Werner, RR Fay, AN Popper (Eds.), Springer Handbook of Auditory Research, Springer-Verlag, New York, pp. 19–57.

Arbogast TL, Mason CR, Kidd Jr G (2002). The effect of spatial separation on informational and energetic masking of speech. J Acoust Soc Am 112 (5 Pt 1): 2086–2098.

Bargones JY, Werner LA, Marean GC (1995). Infant psychometric functions for detection: mechanisms of immature sensitivity. J Acoust Soc Am 98 (1): 99–111.

Barker BA, Newman RS (2004). Listen to your mother! The role of talker familiarity in infant streaming. Cognition 94 (2): B45–B53.

Birnholz JC, Benacerraf BR (1983). The development of human fetal hearing. Science 222 (4623): 516–518.

Blauert J (1997). Spatial Hearing: The Psychophysics of Human Sound Localization, revised edn. MIT Press, Cambridge, Massachusetts.

Bourgeois JP, Goldman-Rakic PS, Rakic P (2000).In: MS Gazzaniga (Ed.), The New Cognitive Neurosciences, 2nd edn. MIT Press, Cambridge, MA, pp. 45–53.

Bregman AS (1990). Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge, MA.

Bronkhorst AW (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. Acustica 86: 117–128.

Bull D, Schneider BA, Trehub SE (1981). The masking of octave-band noise by broad-spectrum noise: a comparison of infant and adult thresholds. Percept Psychophys 30 (2): 101–106.

Buss E, Hall 3rd JW (2009). Effects of masker envelope coherence on intensity discrimination. J Acoust Soc Am 126 (5): 2467–2478.

Cherry EC (1953). Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25: 975–979.

Ching TY, Dillon H (2013). Major findings of the LOCHI study on children at 3 years of age and implications for audiological management. Int J Audiol 2013 (52 Suppl 2): S65–S68.

Clifton RK, Morrongiello BA, Kulig JW et al. (1981). Newborns' orientation toward sound: Possible implications for cortical development. Child Dev 52: 833–838.

Coch D, Sanders LD, Neville HJ (2005). An event-related potential study of selective auditory attention in children and adults. J Cogn Neurosci 17 (4): 605–622.

Culling JF, Hawley ML, Litovsky RY (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. J Acoust Soc Am 116: 1057–1065.

Doupe AJ, Kuhl PK (1999). Annu Rev Neurosci 22: 567–631.

Durlach NI, Mason CR, Shinn-Cunningham BG et al. (2003). Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. J Acoust Soc Am 114 (1): 368–379.

Eby TL, Nadol Jr JB (1986). Postnatal growth of the human temporal bone. Implications for cochlear implants in children. Ann Otol Rhinol Laryngol 95 (4 Pt 1): 356–364.

Ehlers E, Hess C, Litovsky R (2013). Spatial hearing abilities in toddlers with bilateral cochlear implants. Presented at the 12th Annual Early Hearing Detection and Intervention Meeting, Glendale, AZ.

Elliott LL, Connors S, Kille E et al. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. J Acoust Soc Am 66 (1): 12–21.

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. Rev Mod Phys 12: 47–65.

Flexer C (2011). Cochlear implants and neuroplasticity: linking auditory exposure and practice. Cochlear Implants Int 12 (Suppl 1): S19–S21.

Galambos R, Hecox KE (1978). Clinical applications of the auditory brain stem response. Otolaryngol Clin North Am 11 (3): 709–722.

Garadat SN, Litovsky RY (2007). Speech intelligibility in free field: spatial unmasking in preschool children. J Acoust Soc Am 121 (2): 1047–1055.

Gordon KA, Wong DD, Papsin BC (2013). Bilateral input protects the cortex from unilaterally-driven reorganization in children who are deaf. Brain 136 (Pt 5): 1609–1625.

Grieco-Calub T, Litovsky RY (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 TM, Litovsky RY, Werner LA (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 JH, Hall 3rd JW, Gibbs C (1993). Temporal analysis in children. J Speech Hear Res 36 (2): 351–356.

Hall JW, Haggard MP, Fernandes MA (1984). Detection in noise by spectro-temporal pattern analysis. J Acoust Soc Am 76 (1): 50–56.

Hall 3rd JW, Grose JH, Dev MB (1997). Auditory development in complex tasks of comodulation masking release. J Speech Lang Hear Res 40 (4): 946–954.

Hall 3rd JW, Grose JH, Buss E et al. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. Ear Hear 23 (2): 159–165.

Hall 3rd JW, Buss E, Grose JH (2007). The binaural temporal window in adults and children. J Acoust Soc Am 121 (1): 401–410.

Hartley DE, Wright BA, Hogan SC et al. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. J Speech Lang Hear Res 43 (6): 1402–1415.

Hartmann WM, Rakerd B (1989). On the minimum audible angle – a decision theory approach. J Acoust Soc Am 85: 2031–2041.

Hartung K, Trahiotis C (2001). Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. J Acoust Soc Am 110: 1505–1513.

Hawley ML, Litovsky RY, Colburn HS (1999). Speech intelligibility and localization in complex environments. J Acoust Soc Am 105: 3436–3448.

Hawley ML, Litovsky RY, Culling JF (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. J Acoust Soc Am 115: 833–843.

He S, Buss E, Hall 3rd JW (2010). Monaural temporal integration and temporally selective listening in children and adults. J Acoust Soc Am 127 (6): 3643–3653.

Hess CL (2013). Speech discrimination and spatial hearing in toddlers with bilateral cochlar implants. PhD Dissertation, University of Wisconsin-Madison.

Johnstone PM (2006). Informational Masking and Spatial Asymmetry in a "Cocktail Party" Environment. University of Wisonsin-Madison, Unpublished Dissertation.

Jones GL, Litovsky RY (2008). Effects of uncertainty in a cocktail party environment in adults. J Acoust Soc Am 124: 3818–3830.

Kacelnik O, Nodal FR, Parsons CH et al. (2006). Training-induced plasticity of auditory localization in adult mammals. PLoS Biol 4 (4): e71.

Kawano A, Seldon HL, Clark GM et al. (1998). Intracochlear factors contributing to psychophysical percepts following cochlear implantation. Acta Otolaryngol 118: 313–326.

Keefe DH, Abdala C (2011). Distortion-product otoacoustic-emission suppression tuning in human infants and adults

using absorbed sound power. J Acoust Soc Am 129 (4): EL108–EL113.

Keefe DH, Bulen JC, Arehart KH et al. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. J Acoust Soc Am 94 (5): 2617–2638.

Kral A, Tillein J, Hubka P et al. (2009). Spatiotemporal patterns of cortical activity with bilateral cochlear implants in congenital deafness. J Neurosci 29 (3): 811–827.

Kral A, Heid S, Hubka P et al. (2013). Unilateral hearing during development: hemispheric specificity in plastic reorganizations. Front Syst Neurosci 7: 93.

Leake PA, Hradek GT, Snyder RL (1999). Chronic electrical stimulation by a cochlear implant promotes survival of spiral ganglion neurons after neonatal deafness. J Comp Neurol 412 (4): 543–562.

Leibold LJ, Neff DL (2007). Effects of masker-spectral variability and masker fringes in children and adults. J Acoust Soc Am 121 (6): 3666–3676.

Leibold LJ, Werner LA (2006). Effect of masker-frequency variability on the detection performance of infants and adults. J Acoust Soc Am 119 (6): 3960–3970.

Litovsky R (1997). Developmental changes in the precedence effect: estimates of minimal audible angle. J Acoust Soc Am 102: 1739–1745.

Litovsky RY (1998). Physiological studies of the precedence effect in the inferior colliculus of the kitten. J Acoust Soc Am 103: 3139–3152.

Litovsky RY (2005). Speech intelligibility and spatial release from masking in young children. J Acoust Soc Am 117 (5): 3091–3099.

Litovsky RY (2011). Development of binaural and spatial hearing. Springer Handbook of Auditory Research. Springer-Verlag, New York, pp. 163–195.

Litovsky RY, Delgutte B (2002). Neural correlates of the precedence effect in the inferior colliculus: effect of localization cues. J Neurophysiol 87: 976–994.

Litovsky RY, Godar SP (2010). Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. J Acoust Soc Am 128 (4): 1979–1991.

Litovsky R, Macmillan N (1994). Minimum auditory angle for clicks with simulated echoes: effects of azimuth and standard. J Acoust Soc Am 96 (2): 752–758.

Litovsky RY, Madell J (2009). Bilateral cochlear implants in children. In: L Eisenberg (Ed.), Clinical Management of Children with Cochlear Implants, Plural Publishing, San Diego, CA.

Litovsky RY, Yin TC (1998). Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. J Neurophysiol 80 (3): 1285–1301.

Litovsky RY, Colburn HS, Yost WA et al. (1999). The precedence effect. J Acoust Soc Am 106: 1633–1654.

Litovsky RY, Goupell MJ, Godar S et al. (2012). Studies on bilateral cochlear implants at the University of Wisconsin's binaural hearing and speech lab. J Am Acad Audiol 23 (6): 474–494.

Lutfi RA, Kistler DJ, Oh EL et al. (2003). One factor underlies individual differences in auditory informational masking

within and across age groups. Percept Psychophys 65 (3): 396–406.

Macmillan NA, Creelman CD (2004). Detection Theory: A User's Guide. Lawrence Erlbaum, Hillsdale, NJ.

Maxon AB, Hochberg I (1982). Development of psychoacoustic behavior: sensitivity and discrimination. Ear Hear 3 (6): 301–308.

McAdams S, Bertoncini J (1997). Organization and discrimination of repeating sound sequences by newborn infants. J Acoust Soc Am 102 (5 Pt 1): 2945–2953.

Middlebrooks JC, Green DM (1991). Sound localization by human listeners. Annu Rev Psychol 42: 135–159.

Mills A (1958). On the minimum audible angle. J Acoust Soc Am 30: 237–246.

Misurelli SM, Litovsky RY (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry. J Acoust Soc Am 132 (1): 380–391.

Misurelli SM, Bernstein SR, Johnsrude IS et al. (2013). How do cognitive factors interact with speech-in-noise segregation in normal hearing children and adults? J Assoc Res Otolaryngol, Abstracts Feb 16–23.

Moore JM, Thompson G, Thompson M (1975). Auditory localization of infants as a function of reinforcement conditions. J Speech Hear Disord 40: 29–34.

Moore DR, Ferguson MA, Halliday LF et al. (2008). Frequency discrimination in children: perception, learning and attention. Hear Res 238 (1–2): 147–154.

Muir DW, Clifton RK, Clarkson MG (1989). The development of a human auditory localization response: a U-shaped function. Can J Psychol 43 (2): 199–216.

Neff DL, Green DM (1987). Masking produced by spectral uncertainty with multicomponent maskers. Percept Psychophys 41 (5): 409–415.

Newman RS (2005). The cocktail party effect in infants revisited: listening to one's name in noise. Dev Psychol 41 (2): 352–362.

Newman RS, Jusczyk PW (1996). The cocktail party effect in infants. Percept Psychophys 58 (8): 1145–1156.

Niparko JK, Tobey EA, Thal DJ et al. (2010). JAMA 303 (15): 1498–1506.

Nittrouer S, Boothroyd A (1990). Context effects in phoneme and word recognition by young children and older adults. J Acoust Soc Am 87 (6): 2705–2715.

Nodal FR, Kacelnik O, Bajo VM et al. (2010). Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets. J Neurophysiol 103 (3): 1209–1225.

Nodal FR, Bajo VM, King AJ (2012). Plasticity of spatial hearing: behavioural effects of cortical inactivation. J Physiol 590 (Pt 16): 3965–3986.

Northern JL, Downs MP (1984). Hearing in children, Williams & Wilkins, Baltimore.

Nozza RJ, Wagner EF, Crandell MA (1988). Binaural release from masking for a speech sound in infants, preschoolers, and adults. J Speech Hear Res 31 (2): 212–218.

Oh EL, Lutfi RA (1998). Nonmonotonicity Of informational masking. J Acoust Soc Am 104 (6): 3489–3499.

Oh EL, Wightman F, Lutfi RA (2001). Children's detection of pure-tone signals with random multitone maskers. J Acoust Soc Am 109 (6): 2888–2895.

Olsho LW, Koch EG, Halpin CF (1987). Level and age effects in infant frequency discrimination. J Acoust Soc Am 82 (2): 454–464.

Olsho LW, Koch EG, Carter EA et al. (1988). Pure-tone sensitivity of human infants. J Acoust Soc Am 84 (4): 1316–1324.

Plomp R, Mimpen AM (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.

Pujol R, Hilding D (1973). Early synapse formation in vestibular system of fetal guinea pig. Acta Otolaryngol 76 (1): 1–10.

Scollie SD, Ching TY, Seewald RC et al. (2010). Children's speech perception and loudness ratings when fitted with hearing aids using the DSL v.4.1 and the NAL-NL1 prescriptions. Int J Audiol 49 (Suppl 1): S26–S34.

Serpanos YC, Gravel JS (2000). Assessing growth of loudness in children by cross-modality matching. J Am Acad Audiol 11 (4): 190–202.

Shahidullah S, Hepper PG (1994). Frequency discrimination by the fetus. Early Hum Dev 36 (1): 13–26.

Shepherd RK, Hardie NA (2001). Deafness-induced changes in the auditory pathway: implications for cochlear implants. Audiol Neurootol 6 (6): 305–318.

Shepherd RK, McCreery DB (2006). Basis of electrical stimulation of the cochlea and the cochlear nucleus. Adv Otorhinolaryngol 64: 186–205.

Sinnott JM, Aslin RN (1985). Frequency and intensity discrimination in human infants and adults. J Acoust Soc Am 78 (6): 1986–1992.

Song JH, Skoe E, Banai K et al. (2012). Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex 22 (5): 1180–1190.

Spetner NB, Olsho LW (1990). Auditory frequency resolution in human infancy. Child Dev 61 (3): 632–652.

Sussman E, Wong R, Horváth J et al. (2007). The development of the perceptual organization of sound by frequency separation in 5-11-year-old children. Hear Res 225 (1–2): 117–127.

Tharpe AM, Ashmead DH (2001). A longitudinal investigation of infant auditory sensitivity. Am J Audiol 10 (2): 104–112.

Trehub SE, Bull D, Schneider BA (1981). Infants' detection of speech in noise. J Speech Hear Res 24 (2): 202–206.

Trehub SE, Schneider BA, Morrongiello BA et al. (1988). Auditory sensitivity in school-age children. J Exp Child Psychol 46 (2): 273–285.

van de Par S, Kohlrausch A (1999). Dependence of binaural masking level difference on center frequency, masker bandwidth, and interaural parameters. J Acoust Soc Am 106: 1940–1947.

Van Deun L, van Wieringen A, Van den Bogaert T et al. (2009). Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. Ear Hear 30: 178–190.

Veloso K, Hall JW, Grose JH (1990). Frequency selectivity and comodulation masking release in adults and in 6-year-old children. J Speech Hear Res 33: 96–102.

Watson CS, Kelly WJ, Wroton HW (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. J Acoust Soc Am 60 (5): 1176–1186.

Whitfield IC, Cranford J, Ravizza R et al. (1972). Effects of unilateral ablation of auditory cortex in cat on complex sound localization. J Neurophysiol 35 (5): 718–731.

Wilson BS, Dorman MF (2008). Cochlear implants: current designs and future possibilities. J Rehabil Res Dev 45 (5): 695–730.

Wright BA, Bowen RW, Zecker SG (2000). Nonlinguistic perceptual deficits associated with reading and language disorders. Curr Opin Neurobiol 10 (4): 482–486.

Wright BA, Lombardino LJ, King WM et al. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. Nature 387 (6629): 176–178.

Xia J, Brughera A, Colburn HS et al. (2010). Physiological and Psychophysical Modeling of the Precedence Effect. J Assoc Res Otolaryngol 11 (3): 495–513.

Yost WA (1997). The cocktail party problem: Forty years later. In: RH Gilkey, TR Anderson (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments, Lawrence Erlbaum, Hillsdale, NJ, pp. 329–348.

Zurek PM, Durlach NI (1987). Masker-bandwidth dependence in homophasic and antiphasic tone detection. J Acoust Soc Am 81: 459–464.