

CHAPTER 2

*The Infant's Auditory World: Hearing, Speech,
and the Beginnings of Language*

JENNY R. SAFFRAN, JANET F. WERKER, and LYNNE A. WERNER

| | | | |
|---------------------------------------------------------|----|-----------------------------------------------------|----|
| INFANT AUDITION | 59 | Stress and Phonotactic Cues | 80 |
| Development of the Auditory Apparatus: | | Higher-Level Units | 81 |
| Setting the Stage | 59 | LEARNING MECHANISMS | 81 |
| Measuring Auditory Development | 60 | Units for Computations | 83 |
| Frequency Coding | 61 | BUILDING FROM THE INPUT DURING THE | |
| Intensity Coding | 63 | 1ST YEAR | 83 |
| Temporal Coding | 67 | Learning Phonology and Phonotactics | 84 |
| Spatial Resolution | 69 | Word Segmentation | 84 |
| Development of Auditory Scene Analysis | 70 | Beginnings of Word Recognition | 88 |
| Implications for the Development of | | Listening for Meaning | 89 |
| Speech Perception | 71 | Beginnings of Grammar | 91 |
| INFANT SPEECH PERCEPTION AND WORD | | CONCLUSIONS AND FUTURE DIRECTIONS | 92 |
| LEARNING: BEGINNINGS OF LANGUAGE | 72 | Relationship between Auditory Processing and | |
| Emergence of the Field: Phonetic Perception | 72 | Speech Perception | 92 |
| A Preference for Speech | 75 | Constraints on Learning | 93 |
| Perception of the Visible Information in Speech | 76 | Domain Specificity and Species Specificity | 93 |
| Perception of Prosodic Attributes of the | | The Infant's Auditory World | 94 |
| Speech Signal | 77 | REFERENCES | 95 |
| Perception of Other Aspects of the Speech Signal | 78 | | |
| IMPLICIT DISCOVERY OF CUES IN THE | | | |
| INPUT: A DRIVE TO MAKE SENSE OF | | | |
| THE ENVIRONMENT | 80 | | |

The auditory world provides a rich source of information to be acquired by the developing infant. We are born with well-developed auditory systems, capable of gathering a wealth of knowledge even prior to birth. Audi-

tion provides a channel for many important sources of inputs, including a variety of critical environmental sounds such as music and spoken language. For these reasons alone, it has long been of great interest to characterize the nature of this system as it develops.

We thank Dick Aslin, Suzanne Curtin, LouAnn Gerken, Lincoln Gray, Jim Morgan, and Erik Thiessen for their helpful comments on a previous draft. Support for the preparation of this chapter was provided by grants from NIH (R01 HD37466) and NSF (BCS-9983630) to J. R. S., from the Natural Science and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada, The Canada Research Chair Program, and the Human Frontiers Science Program to J. F. W., and from NIH (R01 DC00396, P30 DC04661) to L. A. W.

But studying how infants use their auditory environments can tell us more than just how these developmental processes unfold. In the past decade, studies of infant audition, speech, and the beginnings of language have increasingly begun to bear on central debates in developmental cognitive science and cognitive neuroscience. We are moving beyond such classic questions such as whether speech is special (i.e., subserved by a dedicated neural system that is not shared by other aspects of perception) to begin to study the actual learn-

This chapter is dedicated to the memory of Peter Jusczyk.

ing mechanisms underlying infants' precocious acquisition of the speech sounds of their native language. Similarly, studies of the origins of infants' linguistic knowledge have moved beyond descriptions of *when* infants know about various features of their native language to studies that ask *how* that learning occurred. Increasingly, such behavioral studies are paired with research using psychophysiological methods to study the neural underpinnings of the behavior, experiments using nonhuman animals to probe the species-specificity of the behavior, and studies using materials drawn from other domains to assess the domain-specificity of the behavior.

In this chapter, we review the state of the art in our field, using the ever-increasing interdisciplinarity of research on infant audition, speech perception, and early language acquisition to highlight several themes. One broad theme is the cause of developmental change. Are changes due to maturation of central and/or peripheral neural structures? Or are they due to learning mechanisms, which are continually discovering complex structure in the environment? A related broad theme concerns the nature of these perceptual and learning processes, and the extent to which they are specifically tailored for a single task (e.g., learning about speech) as opposed to available more generally for learning across domains. We will also consider constraints on perception and learning—arising from our perceptual systems, neural structures, species-specific limitations on learning, and domain-specific limitations on learning—that will help to inform our theories of how these processes are related to other aspects of infant development. Finally, we will point to many of the open questions that continue to drive research in this field, and which we hope to see answered in the subsequent edition of this Handbook.

INFANT AUDITION

Most infant auditory research has focused on infants' perception of speech. In subsequent sections of this chapter, we review many of these studies that demonstrate that even newborns are capable of making many phonetic and other speech distinctions. Clearly infants have the auditory capacity to represent some of the critical acoustic features of speech. Little is known, however, about the acoustic information that infants use to make these fine-grained distinctions. Immature

auditory processing will result in imprecise representations of speech as well as other sounds, and hence limit the information that is available to the infant. In this section, we consider whether limitations of auditory processing may serve to constrain early speech perception.

The auditory system is designed to locate and identify sound sources in the environment. Sounds entering the ear are shaped by the structures of the outer ear to optimize detection of relevant sounds and to allow determination of a sound's location in space. The sound is then analyzed into frequency bands by the inner ear. Periodicity, intensity and temporal fluctuations are represented within each band. This code provides the basis of all auditory perception, but the auditory system must calculate some sound characteristics from the basic code. For example, the shape of a sound's spectrum is extracted and differences between the ears are calculated in the auditory brainstem. Once all of this coding and calculation, referred to as primary processing, is completed, however, the system must still determine which frequency bands emanate from a common source, on the basis of commonalities in frequency, periodicity, intensity, temporal fluctuations, location, and spectral shape. The latter stage of processing is known as sound source segregation or auditory scene analysis (Bregman, 1990). Failure to segregate a sound source from the background makes a listener less sensitive to that sound; factors that promote sound source segregation tend to make listeners more sensitive to a sound. These processes are likely to undergo important developmental change during infancy. Finally, it is important to recognize that attention, motivation, memory and other cognitive processes influence auditory scene analysis, and in a very real way, hearing. These effects are described collectively as "processing efficiency," and they also contribute to auditory development.

Development of the Auditory Apparatus: Setting the Stage

In humans, the inner ear begins to function during the second trimester of gestation. If humans are like other mammals, neural responses to sound are possible as soon as the inner ear begins to transduce sound. The consensus is that scalp-recorded auditory evoked potentials and behavioral responses can be observed in fetuses and in preterm infants as young as 28 weeks gestational age.

60 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

The possibility of prenatal hearing has important implications for understanding the effects of early experience on neural development. Fetuses' experience with sound is severely limited by the sound transmission properties of maternal tissue and amniotic fluid, by the conduction of sound to the fetal inner ear, and by immaturities of the inner ear and auditory nervous system (Smith, Gerhardt, Griffiths, & Huang, 2003). Nonetheless, several studies have documented differential fetal responsiveness to sounds of different intensities and frequencies (e.g., Lecanuet, Gramer-Deferre, & Busnel, 1988; Shahidullah & Hepper, 1994). Further, several studies have demonstrated that prenatal experience with sound can influence later auditory responses, at least in the immediate postnatal period. This is most dramatically shown in the preference for mother's voice shown at birth (DeCasper & Fifer, 1980) and in the preference for a story and/or song heard prenatally (DeCasper & Spence, 1986). At the same time, little is known about the importance of prenatal experience with sound for auditory or other aspects of development.

At term, the neonate is believed to have a mature inner ear (but see Abdala, 2001; for example, Bargones & Burns, 1988; Bredberg, 1968). However, the conduction of sound through the external and middle ear to the inner ear is less efficient in neonates than in adults (Keefe, Bulen, Arehart, & Burns, 1993; Keefe, Burns, Bulen, & Campbell, 1994; Keefe et al., 2000). The transmission of information through the auditory neural pathway is slow and inefficient (e.g., Gorga, Kaminski, Beauchaine, Jesteadt, & Neely, 1989; Gorga, Reiland, Beauchaine, Worthington, & Jesteadt, 1987; Ponton, Moore, & Eggermont, 1996). The implications of this pattern of immaturities for postnatal auditory development are discussed in the sections that follow.

Approximately 2 to 3 in 1,000 infants are born with a hearing loss, 1 in 1,000 with a severe to profound hearing loss. However, 20% to 30% of hearing-impaired children develop hearing loss postnatally. These children can be identified, even as neonates, with appropriate hearing screening (Norton et al., 2000), although in the recent past, the average age of identification of hearing-impaired children was 2½ years. Disruption of nearly all aspects of development, but particularly of language development, is typical in hearing-impaired children. Recent evidence suggests, however, that early identification of hearing loss—with intervention beginning prior to 6 months of age—facilitates the development of language skills (signed and/or spoken) within the normal

range in childhood (Yoshinaga-Itano, Sedey, Coulter, & Mehl, 1998).

Measuring Auditory Development

Sounds differ in three basic dimensions—frequency, intensity, and changes in frequency and intensity over time. The auditory system encodes the frequency and intensity of sound and extracts information about temporal variation. In addition, the auditory system calculates additional information about differences between sounds arriving at the two ears, or interaural differences. In psychoacoustics, auditory capacities are approached from two directions. One is to describe the accuracy with which a dimension is coded, or resolution. The other is to describe the function relating the acoustic dimension to its perception. Both approaches have been taken in characterizing hearing during infancy, although the former is more straightforward in a nonverbal subject.

Except as noted, the studies described in this section used one of three varieties of discrimination learning procedures to estimate infants' thresholds for detecting or discriminating between sounds. Each depends upon teaching the infant that when a sound occurs or when a sound changes in some way, a response will be reinforced by the presentation of an interesting audiovisual event. The common reinforcers are mechanical toys and video displays. Two of the procedures teach infants to make a head turn when the appropriate sound occurs. In one variant, infants learn to turn toward the reinforcer (e.g., Berg & Smith, 1983; Nozza & Wilson, 1984). In the other variant, sounds are presented from one of two loudspeakers on a random schedule. The infant learns to turn toward the speaker producing the sound (e.g., Trehub, Schneider, & Edman, 1980). Infants older than about 6 months of age can be successfully tested using either of these procedures. However, younger infants do not make the crisp directional head turns that older infants do. To get around this difficulty, observer-based conditioning procedures capitalize on whatever response the infant makes to the sound (e.g., Tharpe & Ashmead, 2001; Werner, 1995). In this method, which was originally developed to study infants' visual acuity (Teller, 1979), an observer watching the infant knows when a sound may be presented but not whether it was in fact presented. On the basis of the infant's response, the observer must judge whether or not the sound was presented. The infant is reinforced for producing a response that leads to a correct observer judgment. Infants as

young as 1 month of age have been successfully tested using this technique.

Once children are 3 or 4 years old, they can be tested using a variant of adult psychophysical procedures (e.g., Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). Commonly, three intervals are presented to the child, only one of which randomly contains the signal to be processed. The child is asked to choose the interval containing the different sound. The intervals can be presented with cartoon indicators, making the whole procedure more like a video game.

Frequency Coding

Frequency is coded in the auditory system by two mechanisms. The basilar membrane in the inner ear vibrates in response to incoming sound, and because the stiffness of the membrane varies along its length, each position along the membrane responds maximally to a particular frequency. Hair cells are positioned along the length of the basilar membrane. Outer hair cells provide mechanical feedback that results in higher amplitude and more restricted, or sharper, basilar membrane responses to a given frequency. Each inner hair cell transduces basilar membrane motion into a neural response in the auditory nerve fibers that exclusively contact it. Thus activity in a particular auditory nerve fiber indicates the presence of frequencies within a band about a third of an octave wide. The frequency content of any sound, then, is represented in the pattern of activity across auditory nerve fibers innervating different positions along the basilar membrane. This neural representation of sound is referred to as the place code. Because the basilar membrane vibrates at the frequency of stimulation, the action potentials in auditory nerve fibers tend to occur at the frequency of stimulation. Thus, for frequencies below 5,000 Hz, the intervals between action potentials provide another code for frequency.¹ This phenomenon is known as phase locking, and it provides the basis for the temporal code for frequency. The bulk of evidence from adults indicates that both frequency codes are involved

¹Single auditory nerve fibers do not fire on every cycle of a continuing periodic sound; at frequencies above about 1,000 Hz, single nerve fibers cannot fire fast enough to provide a temporal code for frequency. However, each nerve fiber responds at the same phase of the sound and different nerve fibers randomly respond on different cycles. By combining the responses of many auditory nerve fibers responding to the same frequency, a code for frequency can be derived up to 5,000 Hz.

in determining pitch, the perceptual dimension correlated with sound frequency. Further, in adults, it is the processing of sound in the inner ear that limits the representation of complex sounds.

Frequency Resolution

The most common way to assess the resolution of the place code of frequency is to determine the frequencies of competing sounds that interfere with a listener's ability to detect a frequency-specific sound. The phenomenon of one sound's increasing the difficulty of detecting another is called masking. The interfering sound is the masker; the sound to be detected is called the signal or probe. When masking occurs, the threshold ratio of probe to masker intensity required to detect the probe is higher. It has long been known that one sound will only mask another if their frequencies are separated by less than about a third of an octave (there are exceptions to this rule which are discussed below). It is fairly easy to understand, then, that masking occurs when the masker evokes activity in the same auditory nerve fibers that respond to the probe, and that masking provides a method for assessing the quality of the place code for frequency.

Both behavioral and electrophysiological measures indicate that frequency resolution measured using masking is immature at birth, but is mature by about 6 months of age. For example, Spetner and Olsho (1990) showed that a 4,000 or 8,000 Hz tone was masked by a broader range of frequencies for 3-month-olds than for 6-month-olds and adults. Three-month-olds' frequency resolution was mature only at 1,000 Hz, and 6-month-olds demonstrated mature frequency resolution at all frequencies. Studies by Schneider, Morrongiello, and Trehub (1990) and Olsho (1985) confirmed that frequency resolution was mature at 6 months of age. Although a few studies purported to show immature frequency resolution at 4 years of age (Allen, Wightman, Kistler, & Dolan, 1989; Irwin, Stillman, & Schade, 1986), Hall and Grose (1991) subsequently showed that children of this age had mature frequency resolution when thresholds were appropriately measured. Thus, frequency resolution at low frequencies appears to be mature by birth. At high frequencies, frequency resolution becomes adultlike some time between 3 and 6 months.

Lack of development of the auditory nervous system appears to be responsible for early immaturity of frequency resolution. The consensus is that the inner ear mechanisms responsible for frequency resolution are mature at birth (but see Abdala, 2001; for example, Bargones & Burns, 1988; Bredberg, 1968). However, the

62 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

mature frequency resolution established in the inner ear is not faithfully transmitted through the auditory nervous system. Like the studies based on behavioral measures, several studies based on brainstem evoked-potential measures of frequency resolution report maturity at low frequencies, but not at high frequencies for 3-month-old infants (Abdala & Folsom, 1995a, 1995b; Folsom & Wynne, 1987). By 6 months, these measures indicate mature resolution across the frequency range. The parallels between behavioral and neural evoked potentials results suggest a neural basis for the immaturities observed early in infancy.

Frequency Discrimination

Frequency resolution is a measure of the precision of the place code for frequency. Discrimination between sounds on the basis of frequency, however, is accomplished via both the place code and the temporal code for frequency. Despite the fact that 6-month-olds have mature frequency resolution, frequency discrimination remains immature, at least at low frequencies. Olsho (1984) first reported that 6-month-olds needed about twice the frequency change that adults did to detect a change in frequencies below 2,000 Hz. At 1,000 Hz, several studies have estimated that 6-month-old infants can detect a 1.5% to 3% change in frequency, while adults can detect a change of 1% or less (Aslin, 1989; Olsho, 1984; Olsho, Koch, & Halpin, 1987; Olsho, Schoon, Sakai, Turpin, & Sperduto, 1982; Sinnott & Aslin, 1985). At higher frequencies, Olsho (1984) reported that 6-month-olds detected frequency changes as well as adults did. Sinnott and Aslin (1985) and Olsho et al. (1987) reported results generally consistent with that pattern. Olsho et al. also tested 3-month-old infants and found that they performed similarly to 6-month-olds at low frequencies, but that they had higher frequency discrimination thresholds at high frequencies.

Several studies of preschool and school-age children show that pure-tone frequency discrimination is not adultlike until 10 years of age (Jensen & Neff, 1993; Maxon & Hochberg, 1982; Thompson, Cranford, & Hoyer, 1999). Maxon and Hochberg reported that the discrimination of low frequencies was more immature than that of high frequencies at 4 years, consistent with the results of the infant studies. However, these investigators did report some improvement in discrimination of high frequencies between 4 and 12 years. Nonetheless, the greatest changes in high-frequency discrimination appear to occur during the first 6 months of life, while

the greatest changes in low-frequency discrimination appear to occur between 4 and 6 years of age.

The change in high-frequency discrimination between 3 and 6 months is consistent with the improvement in frequency resolution observed at this age. The nature of the prolonged developmental course for low frequency discrimination is less obvious. It has been suggested that adults use the temporal code to represent low frequency tones and the place code to represent high frequency tones (B. C. J. Moore, 1973). One possible explanation for poor low-frequency discrimination by infants and children is that they do not use the temporal code in pure-tone frequency discrimination or that they use the temporal code inefficiently. However, Allen, Jones, and Slaney (1998) reported that, compared to adults, 4-year-olds' detection is more dependent on the periodicity, or "pitchiness" of a tone. The temporal code is the basis of that sound quality (B. C. J. Moore, 1996). In addition 7-year-olds' pure tone frequency discrimination is more affected by decreases in tone duration than is that of adults', which would suggest that children are more dependent on the temporal code (Thompson et al., 1999). However, by 7 years of age, children are also good at low-frequency discrimination (Maxon & Hochberg, 1982). The other possible explanation for poor low-frequency discrimination is that it takes longer to learn to discriminate between low frequencies than high (Demany, 1985; Olsho, Koch, & Carter, 1988). If infants and young children generally take longer to learn a task than adults do, then they might be at a particular disadvantage in learning to discriminate between low frequencies. An analysis of infant frequency discrimination by Olsho, Koch, and Carter (1988) suggests that training effects might account for some, but not all, of the difference between infants and adults in low-frequency discrimination

Perception of Pitch

The relative importance of the temporal and place codes has been debated extensively in the literature on complex pitch perception. A complex tone consists of multiple frequency components, a fundamental frequency and harmonics. The perception of complex pitch is said to be unitary. That is, although the pitch of a complex tone generally matches the pitch of its fundamental, the complex is perceived as having a single pitch, and the higher harmonics contribute to that percept (B. C. J. Moore, 1996). Clarkson and her colleagues have carried out an impressive series of studies of infants' perception of

complex pitch. In many respects, complex pitch perception in 7- to 8-month-olds appears to be adultlike: Infants are able to categorize complexes on the basis of fundamental frequency, even when the fundamental frequency component is missing from the complex (Clarkson & Clifton, 1985; Montgomery & Clarkson, 1997). Infants have difficulty categorizing inharmonic complexes on the basis of pitch, as do adults (Clarkson & Clifton, 1995). However, when only high-frequency harmonics are present, adults are still able to hear the pitch of the missing fundamental, while infants are not (Clarkson & Rogers, 1995). Because periodicity in the waveform of combined high-frequency harmonics provides the basis of this percept in adults, Clarkson and Rogers' result suggests, again, that infants have difficulty using the temporal code in pitch perception.

Intensity Coding

The primary code for intensity in the auditory system is the firing rate of auditory nerve fibers. There are several studies of developing nonhumans that suggest that immature neurons cannot sustain a response over time and that the maximum firing rate achieved by auditory nerve fibers increases with development (Sanes & Walsh, 1998). In humans, evoked potential amplitude increases more slowly with increasing intensity in infants than in adults (Durieux-Smith, Edwards, Picton, & McMurray, 1985; Jiang, Wu, & Zhang, 1990). Further, because the external and middle ear grow during infancy and childhood, the conduction of sound to the inner would be expected to improve with age. Thus, there is reason to believe that intensity processing would undergo postnatal developmental change in humans.

Intensity Resolution

Some sensory processes can be measured by comparing thresholds across masking conditions. In the studies of infant frequency resolution described earlier (Olsho, 1985; Schneider et al., 1990; Spetner & Olsho, 1990), 6-month-olds' thresholds for detecting a tone were always higher than those of adults, but their thresholds changed as the frequency or bandwidth of the masker changed exactly as adults' thresholds did. It is the pattern of change in threshold that indicates resolution. Making it difficult to judge intensity resolution to be immature is the fact that measures of intensity resolution tend to be absolute measures; they do not depend on a comparison of performance across conditions, but on the absolute

level of performance. It is difficult to distinguish intensity coding effects from processing efficiency effects on an absolute measure. Few experiments have been carried out that allow for this distinction, and a major question in this area is the relative contributions of auditory capacities and processing efficiency to age-related changes in thresholds.

Intensity resolution is typically measured psychophysically by finding the smallest intensity change that a listener can detect. When the change detected is from "no sound" to "sound," we say we are measuring absolute sensitivity. When change is detected in an audible sound, we say we are measuring intensity discrimination, increment detection or masking. In the classic intensity discrimination paradigm, a listener hears two or more sounds, and responds to the more intense. In increment detection, the background sound is continuous; the listener responds when the intensity of the background increases. Simultaneous masking is a special case of intensity discrimination or increment detection in that the addition of a signal to the masker is detected as an increase in the intensity of the stimulus. All of these measures largely depend on the same underlying processes, so the expectation is that they will develop along a similar course, with one exception: The "noise" that limits absolute sensitivity is neural and physiological noise. It is not conducted through the external and middle ear. Immaturity of the conductive apparatus will affect the level of a signal played into the ear, but not the level of the background when absolute threshold is measured. In masking, intensity discrimination and increment detection, the conductive apparatus affects both the signal and the background, leaving the signal-to-noise ratio unchanged. Thus, conductive immaturity will be reflected in absolute thresholds, but not the other measures of intensity resolution.

The most commonly measured aspect of intensity processing is the absolute threshold, the intensity of sound that is just detectable in a quiet environment. Some studies have measured absolute thresholds in infants 3 months and younger. Weir (1976, 1979) estimated the behavioral threshold of neonates, based on their spontaneous responses to tones. The thresholds she measured ranged from 68 dB SPL at 250 Hz to 82 dB SPL at 2,000 Hz, approximately 30 and 70 dB higher than adult thresholds at these frequencies, respectively. Ruth, Horner, McCoy, and Chandler (1983) and Kaga and Tanaka (1980) reported behavioral observation audiometry thresholds for 1-month-olds that are similar to

64 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

those reported by Weir for neonates. However, thresholds measured at 1 to 2 months of age using observer-based procedures are quite a bit lower, about 40 to 55 dB SPL, 35 to 45 dB higher than adults' thresholds (Tharpe & Ashmead, 2001; Trehub, Schneider, Thorpe, & Judge, 1991; Werner & Gillenwater, 1990; Werner & Mancl, 1993). The difference between 1-month-olds' and adults' threshold is about 10 dB greater at 500 Hz than at 4,000 Hz (Werner & Gillenwater, 1990). Whether infants' sensitivity actually improves by 25 dB between birth and 1 month is not clear, given the differences in the procedures used to assess threshold.

By 3 months, thresholds appear to improve by about 10 dB at 500 Hz and by nearly 20 dB at 4,000 Hz (Olsho, Koch, Carter, Halpin, & Spetner, 1988). Compared to adults, 3-month-olds are still about 5 dB less sensitive at the higher frequency. A longitudinal, observer-based study of infants' detection thresholds for a broad noise band confirmed an improvement of about 15 dB between 1 and 3 months (Tharpe & Ashmead, 2001). Between 3 and 6 months, very little improvement is observed in the 500-Hz threshold, but a further 15 dB improvement is seen in the 4,000-Hz threshold (Olsho, Koch, Carter, et al., 1988). Tharpe and Ashmead observed about a 15 dB improvement in threshold for a noise band between 3 and 6 months. In the Olsho et al. study, the performance difference between 6-month-olds and adults at 4,000 Hz is about 15 dB, while at 500 Hz it is about 20 dB. There is general consensus that in the vicinity of 1,000 Hz, 6-month-olds are about 15 dB less sensitive than adults (Berg & Smith, 1983; Nozza & Wilson, 1984; Olsho, Koch, Carter, et al., 1988; Ruth et al., 1983; Sinnott, Pisoni, & Aslin, 1983; Tharpe & Ashmead, 2001; Trehub et al., 1980).

More extensive work has documented absolute sensitivity from 6 months to adulthood. Trehub, Schneider, Morrongiello, and Thorpe (1988) measured thresholds for noise bands centered at different frequencies for listeners ranging from 6 months of age through the school years to adulthood. They report that threshold improves by about 25 dB at 400 Hz, about 20 dB at 1,000 Hz, but only 10 dB at 10,000 Hz. Further, the higher the frequency, the earlier adult levels are achieved: 10 years or later at 1,000 Hz, but before 5 years of age at 4,000 and 10,000 Hz.

The ability to detect a change in the intensity of an audible sound is frequently measured by intensity discrimination, that is, by asking the listener to respond to an intensity difference between sounds. Several studies of infants indicate that they are poorer at intensity dis-

crimination than are adults. Few data are available for infants younger than 6 months of age, but there is evidence that newborns respond to an intensity change in a speech sound as small as 6 dB (Tarquinio, Zelazo, & Weiss, 1990). By 7 to 9 months, Sinnott and Aslin (1985) found that infants detected intensity differences of 6 dB between 1,000 Hz tones, while adults could detect differences of about 2 dB. Kopyar (1997) reported that infants of this age detected differences of 9 dB between tones or between broadband noises. Adults detected differences of about 4 dB between tones, but about 3 dB between noises.

Intensity discrimination has not been examined in children between 9 months and 4 years of age. Maxon and Hochberg (1982) tested intensity discrimination of tones in children older than 4 years. They found a steady improvement in the discrimination threshold from about 2 dB at age 4 years to about 1 dB at 12 years, when the level of tones was near 60 dB above the child's absolute threshold. Thus, by 4 years, intensity discrimination appears to be quite good. Only minor improvements occur thereafter, at least for tones well above absolute threshold.

Increment detection matures somewhat earlier than does discrimination between discrete sounds. Several studies have shown that 7- to 9-month-olds can detect 3 to 5 dB increments in a broad noise band (Berg & Boswell, 1998; Kopyar, 1997; Werner & Boike, 2001), under conditions in which adults detect increments of 1 to 2 dB. Schneider, Bull, and Trehub (1988) reported that 12-month-olds could detect 3 dB increments in a continuous broadband noise, while adults could detect increments less than 1 dB. Berg and Boswell (2000) measured increment detection thresholds in 1- and 3-year-old children, for a 2-octave wide noise band centered at 4,000 Hz. Their results for 1-year-olds are similar to those reported by Schneider et al.; 3-year-olds appeared to be adultlike in this task.

The only thresholds for detection of an increment in a tone were reported by Koyar (1997). Infants do relatively worse in detecting tone increments than noise increments, requiring increments of 8 dB, compared to 2 dB for adults. A number of studies have examined the development of detection of a tone or narrow noise band masked by a noise, essentially the detection of an increment in an ongoing sound. Schneider, Trehub, Morrongiello, and Thorpe (1989) estimated masked thresholds for a 1-octave band noise, centered at frequencies ranging from 400 to 10,000 Hz, masked by a broadband noise, in children ranging from 6 months to

10 years of age and adults. At 6 months, the infants' detection threshold was equivalent to a 7 dB increment; adults detected a 1 dB increment.² The age difference was about the same across the frequency range. Relatively large improvements in performance were reported between 6 and 18 months and between 4 years and 8 years. There was little difference between 10-year-olds and adults. Other studies have examined infants' and children's detection of tones in noise (Allen & Wightman, 1994; Bargones, Werner, & Marean, 1995; Berg & Boswell, 1999; Nozza & Wilson, 1984). They report similar results, although it is clear that there is considerable variability across children in this task (e.g., Allen & Wightman, 1994).

Perception of Timbre

Timbre, or sound quality, is determined by the relative amplitudes of the components of a complex sound, and thus involves the comparison of intensities across frequency. The physical dimension associated with timbre is referred to as spectral shape. Vowel perception and sound localization depend on spectral shape processing. A few studies have examined the development of timbre perception. Seven-month-olds can discriminate between sounds of different timbres, complex tones with the same pitch that contain different harmonics (Clarkson, Clifton, & Perris, 1988). Trehub, Endman, and Thorpe (1990) also showed that infants could categorize tonal complexes on the basis of "spectral shape." The sharpness of infants' representation of spectral shape has not been assessed. Allen and Wightman (1992) used a complex sound with a sinusoidal spectral shape to measure children's threshold for detecting changes in spectral shape. They were unable to elicit discrimination between such complexes in 4-year-olds. Five and 7-year-olds performed the task, but only 9-year-olds performed as well as adults. These results suggest that spectral shape, or timbre, discrimination follows a long developmental course. It is not clear that performance on this task generalizes to vowel perception.

Perception of Loudness

A final measure of intensity processing is loudness. In adults, loudness is measured by having listeners match sounds in loudness or by having them rate loudness by

some means. Children as young as 5 years of age are able to rate the loudness of tones numerically and with line length. Moreover, loudness appears to grow with increasing intensity in the same way for children and adults (Bond & Stevens, 1969; Collins & Gescheider, 1989). The evidence on loudness growth in infants is sparse. Leibold and Werner (2002) examined the relationship between intensity and reaction time in 7- to 9-month-olds and adults. Reaction time decreased with increasing sound intensity in both age groups, but the rate of decrease was greater for infants than for adults. This finding suggests that loudness grows more rapidly with increasing intensity in infants, but again, the implications of this finding for early audition are not clear.

In summary, absolute threshold, intensity discrimination, detection of tones masked by noise and spectral shape discrimination all undergo relatively large age-related improvements during infancy and the preschool years. However, adult levels of performance are not reached until 8 or 10 years of age. Interestingly, increment detection in broadband sounds appears to mature earlier, around 3 years of age. Nozza (1995; Nozza & Hensen, 1999) showed that the level at which a noise would just start to mask a tone was 8 dB more intense for infants than for adults. This clever experiment demonstrates that immature thresholds of 8- to 11-month-olds are due largely to changes in sensitivity, rather than performance factors. Several factors are known to contribute to age-related improvements in intensity processing. For example, the frequency response of the infant ear canal changes during infancy. While the adult ear canal conducts sounds best in the range between 2,000 and 5,000 Hz, the infant ear canal conducts higher frequency sounds better (Keefe et al., 1994). Further, the efficiency with which the middle ear conducts sound into the inner ear has been shown to increase over a long age period, from birth to perhaps 10 years of age (Keefe et al., 1993, 2000; Keefe & Levi, 1996; Okabe, Tanaka, Hamada, Miura, & Funai, 1988). The largest improvements occur during the 1st year of life, especially for frequencies over about 1,000 Hz. It has been estimated that the efficiency of the conductive apparatus improves by as much as 20 dB at 3,000 Hz between birth and adulthood, with about half of that improvement occurring during infancy. Age-related improvement in conductive efficiency is smaller at lower frequencies, with about a 5 dB improvement between birth and adulthood (Keefe et al., 1993). Thus, one factor in the development of absolute sensitivity is the development of the conductive apparatus. It is likely that

²If the threshold is expressed as the sound pressure level of the signal added to the masker, 6-month-olds' threshold for detecting a noise band or a tone in noise is 8 to 10 dB higher than adults'.

66 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

much, but not all, of the improvement in high-frequency absolute thresholds during the first 6 months of life is due to improvements in conductive efficiency. The conductive improvement is also reflected in thresholds for evoked neural responses (e.g., Lary, Briassoulis, de Vries, Dubowitz, & Dubowitz, 1985; Sininger & Abdala, 1996; Sininger, Abdala, & Cone-Wesson, 1997).

Development of the inner ear is probably not a contributing factor to postnatal auditory development. Most indications are that the inner ear is adultlike in structure by term birth (Bredberg, 1968; Fujimoto, Yamamoto, Hayabuchi, & Yoshizuka, 1981; Hoshino, 1990; Igarashi & Ishii, 1979, 1980; Igarashi, Yamazaki, & Mitsui, 1978; Lavigne-Rebillard & Bagger-Sjoberg, 1992; Lavigne-Rebillard & Pujol, 1987, 1988, 1990; Nakai, 1970; Pujol & Lavigne-Rebillard, 1992). Otoacoustic emissions³ may have a higher amplitude in infants and a higher stimulus level may be required to elicit emissions, but they are qualitatively adultlike (Bonfils, Avan, Francois, Trotoux, & Narcy, 1992; Bonfils, Francois, et al., 1992; Brown, Sheppard, & Russell, 1994; Burns, Campbell, & Arehart, 1994). Abdala and her colleagues (Abdala, 1998, 2001; Abdala & Chatterjee, 2003) have presented data on otoacoustic emissions from young infants that differ somewhat from that seen in adults. However, at this point it is not clear that these differences could not be accounted for by the reduction in input to the cochlea because of middle ear immaturity.

Werner and her colleagues (Werner, Folsom, & Mancl, 1993, 1994) have reported that the time that it takes for the neural response to sound to travel through the brainstem predicts a 3-month-old's absolute threshold at 4,000 and 8,000 Hz. Thus, another factor in the early development of intensity processing is maturation of the primary auditory nervous system. Anatomical and electrophysiological studies suggest that the auditory brainstem continues to develop throughout infancy (Gorga et al., 1989; J. K. Moore, Guan, & Shi, 1997; J. K. Moore, Perazzo, & Braun, 1995; J. K. Moore, Ponton, Eggermont, Wu, & Huang, 1996; Ponton, Eggermont, Coupland, & Winkelaar, 1992; Ponton et al., 1996). Ponton et al. (1996) provide evidence that age-related change in evoked potentials during infancy results largely from increases in synaptic efficiency. Primary

auditory cortex takes much longer to reach maturity, with anatomical and physiological changes occurring to adolescence (J. K. Moore, 2002; J. K. Moore & Guan, 2001; Ponton et al., 2000). The extent to which these changes in the auditory nervous system are reflected in intensity processing beyond infancy is not known.

Finally, higher-level processes also contribute to age-related improvements in intensity processing. Such processes fall into the category of efficiency, as opposed to resolution, but nonetheless influence sensitivity. Several investigators have examined the age differences in intensity processing, and concluded that only a small portion of the difference in thresholds between infants and adults can be accounted for by simple lapses of attention (Viemeister & Schlauch, 1992; Werner, 1992; Wightman & Allen, 1992). Recall, however, that infants and children tend to do quite well in detecting increments in a broadband noise. Werner and Boike (2001) showed that 7- to 9-month-old infants were more adultlike in their detection of an increment in a broadband noise than they were in detecting a tone masked by a broadband noise. Werner and Boike argued that because infants ultimately achieve the same asymptotic performance in detecting tones and noise, the difference cannot be accounted for by differences in attentiveness to narrow-band and broadband sounds. Bargones and Werner (1994) showed that adults tended to listen selectively for a tone at an expected frequency, with the result that they did not hear tones at unexpected frequencies. Infants, in contrast, detected expected and unexpected frequencies equally well. This suggests that infants listen over a broad band of frequencies, even when they are detecting a narrow-band sound. As a result, more of the background noise will interfere with their detection and their thresholds will be higher (Bargones et al., 1995; Dai, Scharf, & Buus, 1991). Because infants and adults both listen over a broad frequency range when detecting broadband sounds, infants will detect broadband sounds better, relative to adults. There is evidence that 6-year-olds detect expected frequencies better than unexpected frequencies (Greenberg, Bray, & Beasley, 1970). The age at which this ability is acquired has not been established.

To summarize, maturation of the conductive apparatus, of primary auditory pathways, and of listening strategies are important factors in the development of intensity processing. Infants' failure to distribute their attention to the features of complex sound as adults do may have additional implications. Given that infants

³Otoacoustic emissions are sounds that are produced in the ear and transmitted back out into the ear canal. The presence of otoacoustic emissions is a good indicator of normal cochlear function.

have relatively little experience with complex sounds and don't know which features are most important, their broadband approach seems sensible. An interesting question is how this broadband listening strategy influences what infants hear in a natural environment.

Temporal Coding

As Viemeister and Plack (1993) noted "the temporal pattern of spectral changes is, essentially, the information substrate" of speech and other communication signals (p. 116). Temporal resolution is defined as the precision with which a listener can follow rapid changes in the intensity or frequency of sound over time. Adult listeners process sounds through a running temporal window, which averages the input over about 8 ms (B. C. J. Moore, Glasberg, Plack, & Biswas, 1988). However, adults also demonstrate temporal integration: They can integrate or otherwise combine a series of these 8-ms "looks" at a sound over 200 to 300 ms while maintaining the temporal detail with 8 ms resolution (Viemeister, 1996; Viemeister & Wakefield, 1991). As noted, early in development auditory neurons often do not maintain their response to ongoing sound, at least in nonhuman species (Sanes & Walsh, 1998). Moreover, it is well established that auditory evoked responses from infants and children are more susceptible to adaptation, or the effects of prior stimulation, than responses from adults. (Fujikawa & Weber, 1977; Fujita, Hyde, & Alberti, 1991; Jiang et al., 1990, 1991; Klein, Alvarez, & Cowburn, 1992; Lasky, 1984, 1991, 1993, 1997; Lasky & Rupert, 1982; Mora, Exposito, Solis, & Barajas, 1990; Plessinger & Woods, 1987). Adaptation effects decline with age, but may be seen in the auditory brainstem response as late as 3 years of age (Jiang et al., 1991). More frontally generated evoked potentials may exhibit even more dramatic adaptation effects in infants and children (e.g., Mora et al., 1990). Such effects may be reflected in a reduced ability to follow rapidly changing stimuli. How temporal integration might be affected is unclear.

Conclusions about the development of temporal resolution depend strongly on the measure chosen to describe it. Further, the interpretation of age effects that have been demonstrated is not always obvious. A good example is the development of duration discrimination. At least two studies indicated that infants as young as 2 months of age could discriminate a change in the duration of a repeated 200- or 300-ms-long sound of about

20 ms (Jusczyk, Pisoni, Reed, Fernald, & Myers, 1983; Morrongiello & Trehub, 1987), a change on the order of 10%. Morrongiello and Trehub also found that 5- to 6-year-olds could discriminate 15 ms changes, while adults discriminated 10 ms changes in the same conditions. At the same time, two studies of children 4- to 10-years-old, using an oddity task ("Which of these three sounds is different?"), report that 4-year-olds may need at least a 50% change in the duration of a 300 or 400 ms sound to detect the change (Elfenbein, Small, & Davis, 1993; Jensen & Neff, 1993). Elfenbein et al. found that duration discrimination did not reach adult levels of performance until 10 years of age. There are many differences among these studies, most notably in the methods used to estimate a threshold, making it extremely difficult to decide how good duration discrimination really is at a given age.

The earliest studies of temporal resolution examined gap detection in infants and in children. In a gap detection task, the duration of an interruption, or "gap," in a sound is manipulated to find the shortest gap duration that can be detected. Werner et al. (1992) reported that the gap detection thresholds in a continuous noise for 3-, 6-, and 12-month-old infants were an order of magnitude worse than those of adults, around 50 compared to 5 ms. Trehub, Schneider, and Henderson (1995) reported better performance for 6 and 12-month-olds when the gap to be detected was between two short tone pips. However, infants' gap detection thresholds were still around 30 ms. Six-month-olds barely achieved 70% correct for gap durations of 28 and 40 ms. Wightman and his colleagues (1989) reported that gap detection thresholds remained immature at 3.5 years, with the threshold for detecting a gap in a noise band centered at 2,000 Hz at about 10 ms. Trehub et al. and Wightman et al. both reported that gap detection thresholds were mature by 5 years of age. In adults, gap detection performance tends to be better for high-frequency sounds than for low (Eddins & Green, 1995). Both infants and children demonstrate adultlike frequency effects on gap detection (Werner et al., 1992; Wightman et al., 1989).

Trehub et al. (1995) suggested that the reason that infants had somewhat better thresholds for detecting a gap between two short sounds than they did for detecting a gap in a continuous sound was that the continuous sound created excess adaptation in the immature auditory system. If the neural response is low at the onset of the gap, because of adaptation, it will be more difficult to detect the gap. Further, if the immature system takes longer to

68 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

recover from adaptation, the response at the offset of the gap will be reduced. A psychophysical paradigm for measuring adaptation is forward masking: a relatively intense sound, the masker, is presented, quickly followed by a probe sound. If the interval between the initial and following sound is less than about 100 ms, the audibility of the probe is reduced relative to the unmasked condition. The development of forward masked thresholds has been examined in infants and in children. Werner (1999) measured forward masked thresholds for a 1-kHz tone in 3- and 6-month-old infants at masker-probe intervals ranging from 5 to 100 ms. Her results showed that the amount of forward masking decreased as the interval increased, and in the same way, for each age group. The audibility of the probe tone was affected more for the 3-month-olds than it was for older listeners. However, 6-month-olds were more or less adultlike in the amount of forward masking demonstrated at all intervals. Thus, by this measure of temporal resolution, 3-month-olds are immature, but 6-month-olds are not. This conclusion argues against the idea that adaptation effects are responsible for variation in gap detection thresholds of 6-month-olds. Buss, Hall, Grose, and Dev (1999) also report that the amount of forward masking demonstrated by 5- to 11-year-old children is adultlike.

A measure of temporal resolution that has received considerable attention from developmentalists in recent years is backward masking, in which a probe tone is masked by a relatively intense masker that follows it by a short interval (0 to 50 ms). Tallal and her colleagues (Tallal, Miller, Jenkins, & Merzenich, 1997; Tallal & Piercy, 1973, 1974) have long argued that a deficit in auditory temporal resolution is the underlying cause of specific language impairment. Wright et al. (1997) collected psychophysical data from children with language impairment and typically developing children. They found that children with language impairment did not differ from typically developing children in simultaneous masked thresholds. Their thresholds were a little higher in forward masking, but their thresholds were considerable higher in backward masking. This finding was taken to support Tallal's position and has spurred research in this area. It has also spurred an interest in the development of backward masking, which had heretofore not been examined. For example, Hartley, Wright, Hogan, and Moore (2000) reported 6-year-olds' backward masked threshold for a 1,000-Hz tone to be 34 dB higher than adults'. At 10 years, backward masked thresholds were nearly 20 dB higher than those of adults. Even if the absolute threshold of 6-year-olds are

5 dB or so higher than those of adults, that still means that they are exhibiting 30 dB more masking than adults. Other studies confirm that children at this age are more susceptible to backward masking than adults (Buss et al., 1999; Rosen, van der Lely, Adlard, & Manganari, 2000). Werner (2003) has also reported that 7- and 11-month-olds have higher backward masked thresholds than adults, although the age difference in amount of masking is not clear. A recent study, however, suggests that at least among children, the apparent susceptibility to backward masking may not reflect immaturity of temporal resolution. Hartley and Moore (2002) showed that a listener with normal temporal resolution, but poor processing efficiency, will be relatively more susceptible to backward masking than to forward or simultaneous masking. It should be noted that a similar conclusion could be drawn about the nature of the perceptual deficit associated with language impairment.

The gold standard of temporal resolution measures is the temporal modulation transfer function (TMTF, Viemeister, 1979). Listeners are asked to detect amplitude modulation (AM) in a sound. The depth, or amount, of modulation is manipulated to define the threshold for AM detection. AM detection threshold is estimated over a range of modulation frequencies. For adults, the result is a function, the TMTF, with a "low-pass characteristic": AM detection threshold is fairly constant from a 4-Hz modulation rate to about 50 Hz. Beyond 50 Hz modulation rate, the AM detection threshold grows poorer at rate of about 3 dB per doubling of modulation frequency. The modulation frequency at which AM detection begins to deteriorate is taken as the measure of temporal resolution. Hall and Grose (1994) described the TMTF of 4- to 10-year-old children. The AM detection threshold of 4- to 7-year-olds was poorer than that of adults across the range of modulation rate; 9- to 10-year-olds were adultlike in this respect. However, the shape of the TMTF was the same for all ages; AM detection began to deteriorate at about 50 to 60 Hz in all age groups. When the TMTF becomes mature is uncertain. Levi and Werner (1996) reported AM detection thresholds of 3-month-olds, 6-month-olds and adults at two modulation rates, 4 and 64 Hz. The difference between thresholds at the two modulation rates for 3- and 6-month-olds was 3 dB. This difference suggests that infants have an adultlike TMTF and mature temporal resolution.

The development of temporal integration has also been of recent interest. In adults, increasing the duration of a sound by a factor of 10 leads to a little less than 10-

dB decrease in the absolute threshold for that sound. This means that adults are integrating information about the sound over time nearly perfectly. Sound energy cannot be integrated, however, over intervals longer than 200 to 300 ms. Several studies of infants have reported that the maximum interval over which infants integrate sound energy in detection is similar to the adult value (e.g., Berg & Boswell, 1995; Thorpe & Schneider, 1987). However, it was also reported that increasing the duration of sound had a much greater than expected effect on infants' absolute threshold. For example, Thorpe and Schneider found that increasing the duration of a noise band by a factor of 6.3 leads to a 20-dB decrease in 6- to 7-month-olds' absolute threshold. Berg and Boswell argued that infants' temporal integration was mature, but that infants had difficulty detecting short duration sounds (see also Bargones et al., 1995). Maxon and Hochberg (1982) reported temporal integration data for 4- to 10-year-olds. For durations of 50 ms and longer, thresholds decreased with increasing duration at an adultlike rate, and thresholds leveled off between 200 and 400 ms duration. The only difference between children and adults was at quite short durations: increasing the duration from 25 to 50 ms leads to a 7-dB decrease in threshold at 4 years. By 12 years, the decrease is only 5 dB, but still greater than expected in adults. Thus, as children get older they appear to be able to deal with progressively shorter sounds. The nature of the immaturity is not clear; Berg and Boswell suggest that it actually could result from immaturity in the growth of neural response with intensity (Fay & Coombs, 1983) or that the immature auditory system is less able to process onset responses and hence, transient stimuli.

Spatial Resolution

Locating sound sources in space involves several processes including evaluation of spectral shape and intensity, as well as binaural comparisons. Under normal circumstances, spectral shape is the primary cue to position in elevation, while binaural time and intensity differences are the primary cues to position in azimuth (the plane that runs through your ears parallel to the ground).

Development of the ability to use these cues has been well studied in infants via measurements of the minimum audible angle (MAA), the threshold for detecting a change in the position of a sound source. The MAA in azimuth has been shown to decrease from about 27° at 1 month to less than 5° at 18 months (Ashmead, Clifton, & Perris, 1987; Clifton, Morrongiello, Kulig, & Dowd,

1981; Morrongiello, 1988; Morrongiello, Fenwick, & Chance, 1990; Morrongiello, Fenwick, Hillier, & Chance, 1994; Morrongiello & Rocca, 1987a, 1990). The MAA is adultlike, 1° to 2°, in 5-year-olds. The MAA in elevation decreases from a value greater than 16° at 6 to 8 months to about 4° at 18 months, which is comparable to the adult MAA in elevation (Morrongiello & Rocca, 1987b, 1987c). In adults, the MAA in azimuth is generally smaller than that in elevation, because additional, binaural, cues can be used to localize sounds in azimuth. Interestingly, during infancy the MAA in azimuth is similar to that in elevation (Morrongiello & Rocca, 1987b, 1987c). That the MAA is similar in the two dimensions suggests that infants may rely more heavily on spectral shape in sound localization than on binaural differences. Finally, several studies have suggested that infants are sensitive to the changes in sound intensity that signal a change in sound source distance (Clifton, Perris, & Bullinger, 1991; Morrongiello, Hewitt, & Gotowiec, 1991). The accuracy with which infants can judge sound source distance has not been examined.

Humans and other mammals base their judgments of a sound source's location on information that first reaches their ears; they are able to suppress information carried in echoes of the original sound. This effect is known as the precedence effect. It is known that infants also demonstrate this effect (Clifton, Morrongiello, & Dowd, 1984). Interestingly, Litovsky (1997) found that while sound localization is influenced to some degree by the presence of echoes in adults, 5-year-olds are more affected. This difference suggests that while the traditional MAA is mature by 5 years, sound localization in real environments may continue to be refined beyond that age.

The mechanisms underlying the development of sound localization are not completely understood. One obvious change that will influence this ability is the growth of the head and external ear. As the head grows, the size of interaural differences increases (Clifton, Gwiazda, Bauer, Clarkson, & Held, 1988), and discrimination of interaural time differences has been shown to improve during infancy (Ashmead, Davis, Whalen, & Odom, 1991). Ashmead et al. (1991), however, showed that immaturity of interaural discrimination does not appear to be great enough to account for early immaturity of sound localization. One possible explanation for early sound localization immaturity is that infants are more dependent on spectral shape than interaural differences in determining a sound source location. Infants'

70 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

ability to use spectral shape cues will also improve as the external ear grows. Another explanation is that infants can process the cues to sound location adequately, but that they have not yet developed the ability to translate a set of acoustic cues into a precise location in space (Gray, 1992). In other animals, it has been shown that multimodal experience is required to grow such a map of sensory space (e.g., Binns, Withington, & Keating, 1995; King, Hutchings, Moore, & Blakemore, 1988). Further, humans who have only monaural hearing early in life may be able to discriminate interaural differences normally when hearing is restored to the previously deprived ear, but still be unable to locate sounds in space (Wilmington, Gray, & Jahrsdorfer, 1994).

Besides allowing us to localize sounds with precision, interaural differences improve our sensitivity to sound. In the laboratory this improvement in sensitivity is called the masking level difference (MLD): Threshold for a tone presented to both ears is lower if there are interaural differences in the masker presented to the two ears. It appears that infants derive less benefit from interaural differences than adults do, and that by 5 years of age the MLD is adultlike (Hall & Grose, 1990; Nozza, 1987). However, 5-year-olds may still derive less benefit from listening with two ears when the listening situation is complex (Hall & Grose, 1990).

Development of Auditory Scene Analysis

Once the auditory system has analyzed incoming sound and extracted information about its spectral shape, temporal fluctuations and location, it remains to resynthesize the auditory scene. Information in different frequency bands must be grouped according to source on the basis of the initial analysis. Moreover, once the scene has been reconstructed, the listener may choose to attend to one sound source, while ignoring others. The development of these processes has not been studied extensively. A few studies suggest that the process of grouping components on the basis of source, called sound source segregation, is functional in infancy, but it is not clear how accurately or efficiently it operates.

Demany (1982), for example, used repeating tone sequences to study source localization. In one sequence, three of four tones were close in frequency while the fourth was somewhat higher in frequency. Adult listeners perceived this sequence as coming from two sources, one producing three different low frequencies, the other producing the single higher frequency tone. If the order

of the tones in this sequence was reversed, adults had no trouble reporting the change. In another sequence, two pairs of near-frequency tones were repeated. Adults heard this sequence as coming from two sources, each of which produced two alternating tones. When the order of this sequence was reversed, adults had difficulty hearing the change. Demany tested 2- to 4-month-olds' ability to discriminate an order change in these two sequences, using a habituation/dishabituation task wherein looking time was the dependent variable. Infants appeared to discriminate the order change in the first sequence, but not in the other sequence, paralleling adults' perception. This result suggests that infants can organize sounds on the basis of frequency.

Demany's (1982) study has been criticized on methodological grounds. It is possible to discriminate some of the sequences he used from their reversed version on the basis of the frequency contour, even if the sequence is not perceived as two parallel streams. Fassbender (1993) corrected this problem and tested 2- to 5-month-olds on sequences that adults organized on the basis of frequency, amplitude or timbre. Infants discriminated order changes in the sequences as adults did, supporting the idea that infants group sounds at least qualitatively like adults. In addition, McAdams and Bertoncini (1997) tested 3- to 4-day-old infants on sequences that adults segregated on the basis of both location and timbre. Again, infants discriminated order reversals as adults did, although it is not clear whether the sequences were organized by location, by timbre, or by both location and timbre. In this test paradigm, note, listeners are never asked to segregate simultaneously occurring sounds, as most frequently occurs in natural environments. Thus, the conclusions that can be drawn about infants' sound source segregation are currently limited.

Only one study has been conducted bearing on the issue of sound source segregation in children. Sound in different frequency bands that fluctuate over time in the same way tend to be grouped together by adults. In fact, an adult will detect a signal at a lower level if the masker consists of multiple frequency bands with common amplitude modulation than when the masker is a single frequency band centered on the signal or if the masker noise bands have different amplitude fluctuations. This effect is known as comodulation masking release (CMR). Grose, Hall, and Gibbs (1993) first showed that 4-year-old children derived the same benefit from adding off-frequency, comodulated frequency bands as

do adults. Hall, Grose, and Dev (1997) subsequently confirmed this finding in slightly older children. However, Hall et al. also reported that when the masker band centered on the signal frequency and the off-frequency comodulated bands were slightly asynchronous, adults' CMR was reduced, but children's CMR was eliminated or became negative. Thus, it would appear that the basic process of grouping frequency bands on the basis of common temporal fluctuations is functional early in life, but the process is more easily disrupted in children than it is in adults. These findings may have considerable relevance to children's listening in modern, complex sound environments.

Finally, to process sound emanating from one among several sources, listeners must be able to ignore irrelevant sounds. Consider that under normal circumstances, the irrelevant sounds in the environment may vary from moment to moment in unpredictable ways. One of the most intriguing findings in psychoacoustics is that uncertainty about the sounds to be ignored makes it much more difficult for listeners to detect a known sound (e.g., Kidd, Mason, & Arbogast, 2002; Neff & Callaghan, 1988; Neff & Green, 1987; Oh & Lutfi, 1999), even when the sounds to be ignored are distant in frequency from the sound to be detected. Reduction in the audibility of one sound due to the introduction of a second sound that does not interfere with the peripheral processing of the signal sound is called informational masking (Pollack, 1975).

In some respects, infants act as if they are uncertain about an irrelevant sound, even when the irrelevant sound does not change over time. Werner and Bargones (1991) showed that 7- to 9-month-olds' thresholds for detecting a tone increased when a noise band distant in frequency was presented simultaneously. Adults did not demonstrate masking under the same condition. If competing distant-frequency tones of varying frequency are presented with the tone to be detected, infants actually exhibit a lesser increase in threshold than adults do relative to the condition in which the competing tones are constant in frequency (Leibold & Werner, 2003). In a sense, infants are less affected by increased uncertainty because they are more uncertain than adults, at least under some conditions.

By contrast, additional uncertainty appears to have more dramatic effects on older children than it does on adults. Allen and Wightman (1995), for example, found that half of 4- to 8-year-olds could not detect a tone at all when two competing tones varying in frequency were

presented. The average threshold of the children who could perform the task was much higher than that of adults. Oh, Wightman, and Lutfi (2001) reported that preschool children demonstrated about 50 dB more masking than adults on average when a varying distant-frequency, two-tone masker was presented with the tone to be detected. Moreover, Wightman, Callahan, Lutfi, Kistler, and Oh (2003) found that while presenting the varying masker tones to the ear contralateral to the signal ear eliminated such informational masking in adults, this manipulation did little to reduce informational masking in preschool children. Since acoustic factors that increase the listener's ability to perceptually segregate the signal and masker typically reduce informational masking in adults, this finding suggests that children's ability to segregate sound sources is not as robust as that of adults. Stellmack, Willihnganz, Wightman, and Lutfi (1997) quantified the extent to which irrelevant information entered into children's perceptual decisions about intensity, finding that preschool children tended to weight information at different frequencies equally, even when they were asked to attend to a single frequency.

Implications for the Development of Speech Perception

The preceding review of infant audition has several implications for their perception of the complex sounds of human language. In the first 6 months of postnatal life, it is likely that the neural representation of sounds is not as sharp or detailed as it is in adulthood. This representational limitation may in turn limit infants' ability to extract information from those sounds. By 6 months of age, infants probably have adultlike representations of speech and other complex sounds. However, this is not to say that their *perception* of complex sounds is adultlike. It is clear that infants do not attend to the information within complex sounds in the same way that adults do. They do not appear to focus on the spectral or temporal details that are most informative. They identify the spatial location of a sound source rather grossly, and they may also have difficulty segregating speech from competing sounds. Adult caregivers may compensate for these immature processing abilities by exaggerating important details and by speaking to infants in a way that makes their speech stand out from background sounds, as indicated at various points in the following sections. We return to the possible links between early auditory

72 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

processing and resulting effects on speech and language learning at the end of this chapter.

INFANT SPEECH PERCEPTION AND WORD LEARNING: BEGINNINGS OF LANGUAGE

Emergence of the Field: Phonetic Perception

When the first studies of infant speech perception were launched in 1971, a number of studies had been published revealing that adults show categorical perception of speech, but not nonspeech sounds. For example, adults presented with an equal step-size continuum of stimuli spanning two phonetic categories (e.g., a voicing difference between /b/ and /p/ or a place of articulation difference between /b/ and /d/) categorically labeled the first several steps along the continuum as one phoneme (e.g., /b/), and the next several steps as the other (e.g., /p/), with a very sharp boundary in between. Moreover, their labeling performance predicted discrimination. When presented with pairs of stimuli of equal sized differences, adults reliably discriminated only those differences to which they were able to assign different phonetic category labels. This perceptual skill is very important to language processing. There are tremendous variations in the way each individual phoneme is pronounced as a function of the other phonemes around it (/b/ is somewhat different in "bat" than in "boot" due to the coarticulation from the following vowel), as a function of speaking rate, and as a function of the voice quality of the individual speaking. Categorical perception allows listeners to treat these differences as equivalent, and thus to recover the word (and hence the meaning) rapidly when listening to others speak. On the basis of the studies published until 1970, it was believed that categorical perception, and perceptual normalization for speaking rate, vowel context, and so on, was unique to humans and unique to speech versus other types of acoustic signals (see Liberman, Cooper, & Shankweiler, 1967; Repp, 1984).

To explore the ontogeny of this capacity, Eimas and his colleagues (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) published a classic study using the high amplitude sucking method, in which infant habituation and dishabituation are measured via rate of sucking on a nonnutritive pacifier (see also Moffit, 1971). Their results demonstrated that 1- and 4-month-old infants, like English-speaking adults, are better able to dis-

criminate stimuli from the /ba-/pa/ continuum that constitute between, rather than within, category differences (according to adult perceptual performance). Given the difficulty of obtaining labeling data from infants, these findings showing better between than within category discrimination, were taken as evidence that infants also show categorical perception. A year later, Morse (1972) extended this work to show that 2-month-old infants can categorically discriminate /ba/ versus /ga/, but fail on nonspeech counterparts to these syllabic forms. Similarly, Eimas (1974) showed that infants, like adults, only discriminate between stimuli that adults label as instances of different categories, holding acoustical distinctiveness constant. A number of additional studies extended this work to other consonant types (e.g., Eimas, 1975a; Hillenbrand, 1984), to consonants in medial as well as initial position (Jusczyk, Copan, & Thompson, 1978), and even to newborn infants (Bertoncini, Bijeljac-Babic, Blumstein, & Mehler, 1987). Subsequent studies demonstrated that phonetic categories in young infants show many other properties observed in adults. The boundaries between categories are not absolute values, as would be indicated by auditory models. Instead, they are influenced by other articulatory variables, such as speaking rate (Miller, 1987). Infants (Eimas & Miller, 1992), like adults (Whalen & Liberman, 1987) show a phenomenon called "duplex perception" wherein the exact same stimulus can be simultaneously heard as both speech and nonspeech with categorical perception of the speech percept and continuous (no sharp category boundaries) perception of the nonspeech percept.

Similar results were found with vowels. Infants, like adults, show categorical perception of brief (Swoboda, Morse, & Leavitt, 1976; Trehub, 1973) but not more extended, isolated vowels (Swoboda, Kass, Morris, & Leavitt, 1978), and categorize vowels as equivalent even across variations in speaker and gender (Kuhl, 1979). In more recent work, Kuhl and colleagues provided data suggesting that vowel categories are organized around prototypes, with "best" central instances (see Grieser & Kuhl, 1989). These central instances have been described as "magnets" which warp the vowel space (Kuhl, 1991; though see Lotto, Kluender, & Holt, 1998, for an alternative account). A comprehensive review of this sizeable early work on consonants and vowels can be found in Eimas, Miller, and Jusczyk (1987). Taken together, these studies led to the claim that speech perception is special in infants just as it is in adults, and

must therefore reflect the operation of a domain-specific ability.

A number of studies examining cross-language speech perception complement the above studies of native-language speech perception. Languages differ in many properties, including their phoneme inventories. English, for example, contains a contrast between /r/ and /l/ which is lacking in Japanese, but English lacks the retroflex /D/ versus dental /d/ distinction that is used in Hindi and other South Asian languages. A series of cross-language speech perception studies in the 1970s revealed that adults have difficulty perceiving acoustically similar nonnative contrasts, and are constrained to distinguishing only those differences that are used phonemically in the native language (e.g., Lisker & Abramson, 1971; Strange & Jenkins, 1978) whereas young infants discriminate phonetic contrasts whether or not they are used in the language they are learning (Aslin, Pisoni, Hennessy, & Percy, 1981; Lasky, Syrdal-Lasky, & Klein, 1975; Streeter, 1976; Trehub, 1976). To capture this pattern of results, Eimas (1975b) proposed that babies are born with broad-based, universal sensitivities and that lack of listening experience leads to loss of unused initial sensitivities. Aslin and Pisoni (1980) formalized this view in their “universal theory” of speech perception, drawing on the notion of “maintenance” as the perceptual mechanism accounting for cross-linguistic differences (Gottlieb, 1976; Tees, 1976).

However, these comparisons of infants and adults relied on different testing procedures for the two populations, and in most cases, tests in different labs on different contrasts. Werker, Gilbert, Humphrey, and Tees (1981) addressed this problem by comparing English infants aged 6 to 8 months, English adults, and Hindi adults on their ability to discriminate both an English and two (non-English) Hindi consonant contrasts, using precisely the same methodology—the conditioned head-turn procedure—with the three groups. Their results confirmed the developmental change. All three groups discriminated the English ba-da contrast, but only the Hindi adults and the English infants discriminated the two Hindi distinctions.

Werker and her colleagues subsequently completed a series of studies designed to identify the age at which the change from “universal” to language-specific phonetic perception might occur (Werker & Tees, 1983), and found important changes occurring across the 1st year of life. At 6 to 8 months of age, English-learning in-

fants successfully discriminate the Hindi retroflex-dental distinction and another non-English (Interior Salish, Nthlakampx glottalized velar versus uvular) distinction, but by 10 to 12 months of age English infants were no longer discriminating non-English distinctions (Werker & Tees, 1984). Confirming that the change was one of maintenance via language-specific listening exposure, and not simply a general decline at 10 to 12 months for all difficult phonetic contrasts, Werker and Tees (1984) showed that infants of the same age (10 to 12 months) raised with Hindi or Nthlakampx did successfully discriminate the contrasts from their native languages.

In the years since this initial work, there have been a number of replications and extensions of this finding. Several studies have confirmed an effect of listening experience on the phonetic differences infants can discriminate by the end of the 1st year of life. A decline in performance on nonnative contrasts was the common pattern in the first wave of studies (Best, McRoberts, Lafleur, & Silver-Isenstadt, 1995; Pegg & Werker, 1997; Tsushima et al., 1994; Werker & Lalonde, 1988; Werker & Tees, 1984). Moreover, the basic pattern of findings from these behavioral studies has been replicated by recording event-related potentials (Cheour et al., 1998; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). Werker showed in her earliest work that this decline likely involves a reorganization of attention rather than a loss of basic discriminatory capacity (Werker & Logan, 1985).

In the past several years, some revealing exceptions to this pattern of findings have appeared. It is now known that vowel perception likely reorganizes at a somewhat younger age than consonant perception (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994), and that acoustically quite distinct contrasts that lie outside the phonological space of the native language (e.g., click contrasts) may remain discriminable even without listening experience (Best, McRoberts, & Sithole, 1988). Indeed, there are differences across nonnative contrasts, with some showing the pattern of decline noted above, and others remaining discriminable (e.g., Best & McRoberts, 2003; Polka & Bohn, 1996), with one influential model suggesting that the assimilability to the native language phonology is the best predictor of maintenance versus decline (Best, 1994). Moreover, experience not only maintains native distinctions, but also seems to improve the sharpness of the native categories (Kuhl, Tsao, Liu, Zhang, & de Boer, 2001; Polka, Colantonio, & Sundara, 2001). This

74 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

observation has led to a reanalysis of the original Eimas (1974) claim of universal phonetic sensitivities from birth, with experience playing primarily a maintenance role. Now it appears that although substantial organization may be evident from birth, learning also plays a role (Kuhl, 2000; Werker & Curtin, 2005). As we will later address further, an exciting new development is attempting to ascertain just how powerful learning might be.

In addition to the empirical evidence suggesting innate biases in phonetic perception, considerable research supported the notion that speech might be special. Early studies suggested that while speech is perceived categorically, listeners typically perceive nonspeech analogues in a more continuous fashion (e.g., Mattingly, Liberman, Syrdal, & Halwes, 1971) and utilize specialized areas or structures in the left hemisphere when engaged in phonetic discrimination tasks (Phillips, Pellathy, & Marantz, 1999; Studdert-Kennedy & Shankweiler, 1970). Similar types of findings were revealed in infant studies (e.g., Eimas et al., 1971). Dichotic tasks with infants showed a significant right ear (LH) advantage in phonetic discrimination tasks by 3 months of age (Glanville, Levenson, & Best, 1977) and possibly earlier (Bertoncini et al., 1989). Even in these initial reports, however, some results did not support the notion of a right-ear/left-hemisphere advantage for phonetic discrimination in young infants (see Best, Hoffman, & Glanville, 1982; Vargha-Khadem & Corballis, 1979, for contradictory studies).

The use of electrophysiology, with the event related potential (ERP) as the dependent variable, has helped clarify the early neuropsychological work. For example, Dehaene-Lambertz and Baillet (1998), using ERP, found brain areas that show activation to a change in phonetic category, but not to an equal sized change within a phonetic category in 3-month-old infants. More recently, they have shown the same pattern of findings when multiple voices are used, showing that the infant brain can extract phonetic categories across variations in speakers (Dehaene-Lambertz & Pena, 2001). ERP studies also consistently reveal asymmetries in phonetic discrimination tasks, but the pattern of asymmetries appears to vary with stimulus type and infant age. Some studies reveal a left hemisphere (LH) advantage for stop consonant discrimination (Dehaene-Lambertz & Baillet, 1998; Dehaene-Lambertz & Dehaene, 1994; Molfese & Molfese, 1979, 1980, 1985) whereas others indicate bilateral responses at birth, with the emergence of right hemisphere (RH) dominance at 3 months of age

(e.g., Novak, Kurtzberg, Kreuzer, & Vaughan, 1989). Molfese consistently found the LH advantage for place-of-articulation differences reported above, but a stronger RH ERP response to a change in voicing (Molfese, Burger-Judish, & Hans, 1991; Molfese & Molfese, 1979). The pattern seen with vowels also suggests asymmetrical processing, in this case favoring the RH (Cheour-Luhtanen et al., 1995). The findings of early sensitivity to potential phonetic distinctions and the possibility of specialized neural systems subserving this discrimination strengthened the notion that "speech is special" and computed by dedicated neural systems.

As is the case with any complex research endeavor, not all the data fit the pattern so nicely. Shortly after the first studies were published revealing categorical-like perception for speech sounds, similar studies were published showing that both adults (Pisoni, 1977) and infants (Jusczyk, Pisoni, Walley, & Murray, 1980) perceive some nonspeech sounds categorically. Moreover, nonhuman animals seem to show similar category boundaries to human infants. Chinchillas (Kuhl & Miller, 1978; Kuhl & Paden, 1983) show categorical perception for both voicing (e.g., /pa/-/ba/) and place (e.g., /ba/-/da/) continua, and several other animal species can also discriminate between consonants (Morse & Snowdon, 1975; Waters & Wilson, 1976). Japanese quail show perceptual constancy of consonant categories across variations in vowels (Kluender, Diehl, & Kileen, 1987), and budgies similarly discriminate consonants (Dooling, Best, & Brown, 1995). Similar findings have been reported for vowels. Monkeys and even cats discriminate /i/ from /u/ (Dewson, 1964), and studies with old world monkeys suggest the same pattern of vowel perception as seen in humans, with excellent discrimination of distinct vowels and more confusion of close vowels such as /E/ (as in bet) and /ae/ (as in bat; Sinnott, 1989). It was initially thought that only humans (adults and infants) show the prototype magnet effect (see Kuhl, 1991), but even rats and birds show a warping of their perceptual space to reveal a prototype organization following brief exposures to vowel categories (Kluender, Lotto, Holt, & Bloedel, 1998; Pons, in press). The animal work raises the strong possibility that speech perception is not necessarily a specialized human capacity, but perhaps instead reflects perceptual biases that are common at least across primates, and perhaps beyond.

To summarize, the first generation of infant speech perception work took as its starting point the work on

phonetic perception in adults, and was designed to assess whether infants showed the same types of responses to phonetic differences as do adults, and if so, whether they used the same underlying neural mechanisms. This research led to increasingly sophisticated methods and techniques, and to studies with both human and nonhuman animals. The findings greatly enriched our understanding of the development of speech perception, and the explanations offered to explain these findings provided rich theoretical fodder for subsequent work. However, one of the insights guiding research for the past several years has been the realization that there is much more to speech than just phonetic categories, and that infants may be sensitive to many other characteristics of the speech around them.

A Preference for Speech

One of the first questions one might ask is whether infants' perceptual systems help them to separate speech from other types of acoustic signals in the environment. An early appearing preference for speech would help infants orient to just those signals in the environment which are essential for language acquisition. Although it is widely believed that infants prefer to listen to speech over other sounds from the first moments of life, there are actually very few data that specifically address this question. Indeed, until very recently the studies upon which this widely held belief was founded were not actually designed to test infants' preference for speech over nonspeech. For example, one study that is widely cited as showing a neonatal preference for speech over nonspeech showed that 4- to 5-month-old infants look longer to a target when it is paired with continuous female speech than when it is paired with white noise (Colombo & Bundy, 1981). Today no one would accept the aversive sound of white noise as an appropriate control for human speech, but in fact Colombo and Bundy had not designed the experiment to test for a preference for speech. Rather, they were attempting to develop a method for assessing infants' responsiveness to different types of speech sounds. Moreover, what is not noted is that 2-month-old infants in Colombo and Bundy's (1981) study, in contrast to older infants, did not respond differently to speech and white noise. The only other early study directly assessing a listening preference for speech was one by Glenn, Cunningham, and Joyce (1981) in which 9-month-old infants pulled a lever more frequently to listen to a female singing a song in

comparison with three solo musical instruments playing the same tune.

More recently, a set of studies has examined infants' preference for acoustic stimuli that have the structural properties of isolated syllables of human speech in comparison to carefully matched nonspeech tokens. Vouloumanos, Kiehl, Werker, and Liddle (2001) used complex nonspeech analogues modeled on sine-wave analogues of speech (Remez, Rubin, Pisoni, & Carrell, 1981). The speech stimuli consisted of the syllable "lif" repeated several times in the high pitch, highly modulated speech that parents use when speaking to their infants. The nonspeech counterparts replaced the fundamental frequency and three most intense higher order frequency components (formants) with a sine wave that tracked their changes across time. Thus, in contrast to the earlier studies investigating preference for speech over nonspeech, these stimuli were carefully matched for duration, timing, fundamental frequency, and area in the spectrum in which information was presented. However, human vocal tracts cannot produce the sine wave nonspeech stimuli.

In the first set of studies, Vouloumanos and Werker (2004) used a sequential preferential looking procedure (e.g., Cooper & Aslin, 1990) to test the listening preferences of 2- to 6-month-old infants. The infants preferred the speech over the complex nonspeech analogues, listening longer on the alternating trials during which speech versus nonspeech was presented. In the second set of studies, newborn infants were tested with these same stimuli, with HAS (high amplitude sucking) as the dependent variable. Like their older counterparts, the newborn infants chose to deliver more HA sucks on the alternating minutes in which speech versus complex nonspeech was presented (Vouloumanos & Werker, 2002). To attempt to rule out a role of experience in eliciting this preference, Vouloumanos created stimuli that would sound like those available to the fetus by filtering them using the filtering characteristics of the uterine wall. Neonates treated the filtered speech and nonspeech as equivalent, even though they discriminate the nonfiltered counterparts. This strengthens the possibility that the preference seen in the newborn is not a direct result of prenatal listening experience with human speech, and argues instead for an evolutionarily given perceptual predisposition for sounds that have the structural characteristics of those which could be produced by a human vocal tract. Moreover, the preference for communicative signals extends beyond spoken language.

76 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

In a recent study, Krentz and Corrina (2005) has shown that hearing infants show a preference for watching sign language over carefully matched nonlinguistic gestures. Taken together, these studies suggest a broad-based perceptual bias for communicative signals.

These results are corroborated by studies using neuroimaging techniques. In an event-related fMRI study completed with adults, Vouloumanos et al. (2001) found that the typical speech areas in the left hemisphere of the temporal cortex are more activated by a change to the speech than a change to the complex nonspeech stimuli described above. This finding complements many other studies with adults showing greater activation of specialized brain areas in the left hemisphere in response to speech than to other types of sounds (e.g., Benson et al., 2001; Binder et al., 1997; Fiez et al., 1995; Price et al., 1996; Zatorre, Evans, Meyer, & Gjedde, 1992; but see also Binder et al., 2000; Zatorre, Meyer, Gjedde, & Evans, 1996).

To date, only two studies have used imaging techniques to determine if the infant brain responds differently to speech versus nonspeech. Both of these studies have contrasted the perception of forward versus backwards speech (for related adult studies, see, Dehaene et al., 1997; Wong, Mihamoti, Pisoni, Sehgal, & Hutchins, 1999). Using both optical topography (Pena et al., 2003), and fMRI (Dehaene-Lambertz, Dehaene, & Hertz-Panier, 2002), speech stimuli elicited greater activation in the infants' LH than the RH. In the Pena et al. (2003) study neonates were tested, and the increased activation was in the classic language areas over the temporal lobe. The Dehaene, Dehaene, and Hertz-Panier (2002) study tested 3-month-olds, eliciting bilateral activation to both the forwards and backwards speech over the temporal lobes, with greater LH activation more posteriorly. Further experimentation with nonhuman primates is necessary to determine whether these early perceptual and neural markers of human speech perception are specific to humans or are instead part of our shared evolutionary history.

Perception of the Visible Information in Speech

Speech perception involves not only the acoustic signal, but also visible articulatory information. The best known example of this is the McGurk effect (McGurk & MacDonald, 1976). When watching a speaker produce the syllable /ga/ while listening to /ba/, adults typically report perceiving an unambiguous /da/ or /tha/, a syllable

that combines features of both the heard and seen stimulus. This effect is robust across many testing conditions and languages (see Green, 1998 for a review) and has been interpreted as part of our endowment for phonetic perception. Yet, there is also evidence of learning. The McGurk effect is stronger in adults than in children (Hockley & Polka, 1994; MacDonald & McGurk, 1978; Massaro, Thompson, Barron, & Laren, 1986), is reduced further in children who have difficulty articulating (Desjardins, Rogers, & Werker, 1997), and shows the same kind of language-specific influences as is seen in the perception of audible speech with nonnative "visemes" assimilated to those used in the native language (Massaro, Cohen, & Smeele, 1995; Werker, Frost, & McGurk, 1992).

Two kinds of studies have explored whether the visible information in speech is available prior to learning. In one, infants are presented with side-by-side displays of two faces articulating two different syllables. An acoustic syllable that matches the syllable being articulated by one of the faces is then presented at mid-line, and the amount of time the infant looks at each face is recorded. Using this method, Kuhl and Meltzoff (1982) showed that infants of 4.5 months look preferentially at the face articulating the heard vowel sound (/a/ versus /i/). This finding has since been replicated with other vowels (Kuhl & Meltzoff, 1988), with male as well as female faces and voices (Patterson & Werker, 1999), with disyllables (e.g., *mama*, *lulu*; MacKain, Studdert-Kennedy, Speiker, & Stern, 1983), and with a high amplitude sucking method (Walton & Bower, 1993). Moreover, these young infants often display mouth movements themselves that correspond to the concordant bimodal display (Kuhl & Meltzoff, 1988; Patterson & Werker, 1999; 2002), suggesting connections including not only the visual and auditory perceptual modalities, but also articulatory processes. More recently it has been shown that the matching effect is equally robust in 2-month-old infants (Patterson & Werker, 2003). The precocity of this matching ability is particularly striking when compared to other types of biologically important information. Infants do not match gender in the face and voice until 7 to 9 months of age (Walker-Andrews, Bahrck, Raglioni, & Diaz, 1991) even when they are tested using precisely the same stimuli for which they show vowel matching (Patterson & Werker, 2002).

Evidence for the McGurk effect itself is less convincing in the infancy period. Although there are reports of infants' percepts showing the same kind of

“fusion” or “visual capture” as seen by adults when mismatched auditory and visual stimuli are presented (Burnham & Dodd, 2004; Desjardins & Werker, 2004; Rosenblum, Schmuckler, & Johnson, 1997), the effect is not nearly as strong or as consistent as that seen in adults. Taken together, these studies suggest that the infant may be endowed from an early age with a perceptual system which is sensitive to both heard and seen features of phonetic segments, but that this system is perfected and tuned through experience listening to and articulating speech. A recent study suggests that we are not the only primate species to use both heard and seen information in perceiving communicative stimuli (Ghazanfar & Logothetis, 2003), suggesting that the intermodality of speech might be deeply engrained in our evolutionary heritage.

Perception of Prosodic Attributes of the Speech Signal

One of the fundamental characteristics of human languages is their prosody—the musical aspects of speech, including their rhythm and intonation. Languages have classically been categorized according to their predominant rhythmic properties into three major types: stress-timed, syllable-timed, and mora-timed (Abercrombie, 1967; Pike, 1945). Stress-timed languages, like English and Dutch, tend to alternate between strong and weak syllables, and the strong syllables are roughly equally spaced in time, thus the term *stress-timed*. Languages like Spanish and Italian, however, use the syllable as the basic unit of timing; syllables are similarly stressed and roughly equally spaced in time. Finally, languages like Japanese are timed-based on the mora, a rhythmic unit roughly corresponding (in English) to a consonant followed by a short vowel (“the” contains one mora, while “thee” contains two). This nomenclature has been refined and quantified by two key properties: percent vowel per syllable and the variability in the consonant (Ramus, Nespor, & Mehler, 1999). These rhythmic properties of the language influence adults’ processing of speech; speakers of different languages employ different units as their primary unit of segmentation. Speakers of syllable-timed languages (e.g., French, Spanish, Catalan, & Portuguese) show a processing advantage for the syllable (e.g., Mehler, Dommergues, Frauenfelder, & Segui, 1981; Morais, Content, Cary, Mehler, & Segui, 1989; Sebastián-Gallés, Dupoux, Segui, & Mehler, 1992), speakers of stress-timed languages such as En-

glish and Dutch show greater access to the phoneme (Cutler, Mehler, Norris, & Segui, 1986; Cutler & Norris, 1988; Vroomen, van Zon, & de Gelder, 1996), and Japanese adults use the mora as the primary unit of segmentation (Otake, Hatano, Cutler, & Mehler, 1993). These differences not only describe the surface properties of languages, but may also provide cues to the underlying syntactic structure, that is, the head direction, of the language (Nespor, Guasti, & Christophe, 1996).

Human infants are sensitive to rhythmical differences from birth. Since the classic study by Demany, McKenzie, and Vurpilot (1977) showing that 2- to 3-month-old infants can discriminate tones based on their rhythmical sequences, there is growing evidence that young infants can use rhythm to detect the timing characteristics of speech (Fowler, Smith, & Tassinari, 1986). Indeed, there is now considerable evidence that infants utilize rhythmical characteristics to discriminate one language from another at birth. In a seminal experiment, Mehler et al. (1988) demonstrated that newborns can discriminate French from Russian produced by a single bilingual speaker—as long as one of these languages is the infant’s native language. Interestingly, at least part of this discrimination is based on prosodic characteristics of speech. The same results were obtained using low-pass filtered speech, in which the phonetic characteristics were removed while retaining the prosodic rhythm of the speech samples. Given that these are the same speech characteristics maintained in the uterine environment, an initial interpretation of these data was that newborn speech capacities are influenced by prenatal maternal exposure. A subsequent reanalysis of the initial data, however, revealed that discrimination was not limited to the native versus an unfamiliar language, but instead was evident as well when the French infants listened to English versus Italian, two unfamiliar languages (Mehler & Christophe, 1995). This weakened the interpretation that it is prenatal experience that leads to language discrimination.

Indeed, there is now abundant evidence that infants use such rhythmical differences to discriminate among a wide variety of languages, suggesting that rhythm may be prioritized in infants’ early speech representations. Consistent with this hypothesis, infants’ discriminations are predictable based on languages’ membership in rhythmic classes. For example, newborns and 2-month-olds can discriminate between two languages from two different classes, but not two languages from the same class (Mehler et al., 1988; Moon, Cooper, & Fifer, 1993;

78 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

Nazzi, Bertoncini, & Mehler, 1998). This discrimination ability is present for forward speech but not backwards speech, which disrupts rhythmic cues. Nonhuman primates show this same pattern of performance, suggesting that processing of the rhythmic characteristics of speech is not specific to humans (Ramus, Hauser, Miller, Morris, & Mehler, 2000). Only by 5 months of age does experience with one's native language allow infants to discriminate it from other languages in the same class (e.g., Nazzi, Jusczyk, & Johnson, 2000). There is thus an interplay between an inherent bias to attend to rhythmic distinctions in languages and the learning processes required to distinguish one's own language from others. Even by 5 months of age, infants can't distinguish between two unfamiliar languages from the same rhythmic class.

Languages from different rhythmical classes differ not only in their timing characteristics, but also in their intonation contours. And, in some cases, this information is also adequate to discriminate two languages that have been low pass filtered at 400 Hz (which removes phonetic content while preserving intonation cues), including English versus Japanese (Ramus & Mehler, 1999). As indicated earlier, young infants are highly sensitive to differences in pitch. Moreover, they are able to use intonation to discriminate vowels (Bull, Eilers, & Oller, 1984; Karzon & Nicholas, 1989), and even to distinguish lists of words differing in pitch contour (Nazzi, Floccia, & Bertoncini, 1998). The preference for infant directed speech (Cooper & Aslin, 1990; Fernald, 1984; Werker & McLeod, 1989) may be explained, at least in part, by this exquisite sensitivity to fundamental frequency (see Colombo & Horowitz, 1986).

It is thus of interest to determine whether infants' ability to discriminate languages is a function of their perception of rhythm, intonation, or both cues in tandem. To do so, Ramus and Mehler (1999) resynthesized natural speech, preserving the rhythm while holding intonation constant. With these stimuli, French newborns were still able to discriminate languages from two different rhythmic classes (Ramus, 2002; Ramus et al., 2000), although levels of discrimination were somewhat attenuated. These results confirm that while infants may be using intonation to boost performance, rhythm alone is sufficient to distinguish one family of languages from another.

These early language discrimination abilities may be particularly useful in bilingual environments. Infants exposed to multiple languages may use these rhythmic distinctions to segregate the input into the different lan-

guages they hear. Such a process would likely facilitate successful bilingual language acquisition by alerting infants to the fact that their input is drawn not from one but two language systems. In the absence of this information, infants may not have a way to determine that they are hearing multiple languages, leading to potential confusions during the learning process.

To begin to address these issues, Bosch and Sebastián-Gallés (2001) assessed the language recognition abilities of 4-month-olds learning both Spanish and Catalan. Importantly, these two Romance languages belong to the same rhythmic category, which should make them quite difficult to discriminate. Nevertheless, these bilingual-to-be infants were able to discriminate between the two languages present in their home environments. These results suggest the availability of an early capacity to distinguish languages given simultaneous bilingual exposure, potentially based on the presence of vowel reduction (since rhythmic cues do not distinguish the two languages). More recently, it has been shown that bilingual-exposed infants process languages differently even from birth (see Werker, Weikum, & Yoshida, *in press*). When tested on their preference for English over a rhythmically distinct language, Tagalog, English-exposed newborns showed a robust preference for filtered English speech over filtered Tagalog speech. However, newborns exposed to both English and Tagalog prenatally did not show this preference, and indeed, chose to listen equally to filtered English and Tagalog. Contrary to hypotheses suggesting that lexical knowledge is needed to engage in language differentiation (e.g., Genesee, 1989), the basic capacities for language differentiation may be in place well prior to the onset of spoken language. However, these results leave open the question of whether infants actually represent the two languages as separate systems, as opposed to discriminable components of a single system.

Perception of Other Aspects of the Speech Signal

In addition to their sensitivity to speech itself, to phonetic segments, to visual speech, and to rhythm and intonation, even the youngest infants also show impressive sensitivities to many other types of information carried by the speech signal. They are sensitive to some kinds of within phonetic category variation. At 3 to 4 months, infants show graded perception of VOT (Miller & Eimas, 1996), and at 6 months can discriminate within-category differences from along the VOT continuum if tested in sufficiently sensitive tasks (McMurray &

Aslin, 2005). Moreover, infants of 6 to 8 months (but not 10 to 12 months) can even treat multiple instances of the voiced, unaspirated [d] versus voiceless, unaspirated [t] (created by removing the “s” from /sta/) as two separate categories, even though adult English speakers treat both of these syllables as equivalently acceptable instances of the phoneme /d/ (Pegg & Werker, 1997).

Sensitivity to within category phonetic variation is necessary in some language processing tasks (see Werker & Curtin, 2005). One illustration comes from the work on allophone discrimination. Allophones are different phonetic realizations of the same phoneme with precise phonetic characteristics that vary depending on their position in words. By 2 months of age, infants can detect the allophonic difference between the unaspirated, unreleased /t/ in “*night rate*” from the aspirated, released, partially retroflex /t/ in “*nitrate*” (Hohne & Jusczyk, 1994), a sensitivity which will ultimately be useful in word segmentation, as discussed next.

Infants are also sensitive to syllable form. Newborn French infants can “count” syllables, discriminating lists of bi- versus trisyllabic words even when the words are modified to have the same overall duration (Bijeljic-Babic, Bertoncini, & Mehler, 1993). They show better discrimination of stimuli that correspond to “good” syllable forms—those with a vocalic nucleus (/tap/ versus /pat/), in comparison to /tsp/ versus /pst/ (Bertoncini & Mehler, 1981). Sensitivity to rhyme (Hayes, Slater, & Brown, 2000), alliteration (Jusczyk, Goodman, & Baumann, 1999), and full syllable repetition (Jusczyk, Goodman, et al., 1999) have all also been demonstrated in infants from 7 to 9 months of age.

One such auditory sensitivity may be of use in the acquisition of grammatically relevant knowledge. Infants are astonishingly sensitive to the acoustic and phonological cues that distinguish grammatical classes. Just as languages differ in their phoneme inventories and in their rhythmical characteristics, they also differ in the number and kinds of grammatical categories words might belong to. For example, while English has prepositions, Chinese has postpositions. The languages of the world seem to all share a fundamental distinction between open-class (lexical words, such as nouns, verbs, adjectives, etc.) and closed-class (grammatical words, such as determiners, prepositions, etc.) categories. These classes of words can be distinguished on the basis of acoustic and phonological cues such as syllable complexity, syllable number, duration, loudness, and presence of reduced vowels (see Kelly, 1992). These differences are magnified in the speech directed to infants,

and are evident in maternal speech across typologically distinct languages (Morgan, Shi, & Allopenna, 1996; Shi, Morgan, & Allopenna, 1998).

In a recent series of studies, Shi and colleagues showed that infants become increasingly able to use these cues across development. Newborn infants categorically discriminate content from function words, even when the words are equated for volume and number of syllables (Shi, Werker, & Morgan, 1999). Specific prenatal listening experience cannot account for this discrimination capacity; the same pattern of results emerges when the items are drawn from an unfamiliar language. By 6 months of age, infants prefer to listen to the content words (Shi & Werker, 2001). This preference cannot be accounted for by specific knowledge of highly familiar items because, again, it is seen even when infants are tested on words from an unfamiliar language (Shi & Werker, 2003). Thus the phenomenon must be based on a developing preference for items with the acoustic and phonological patterns seen in content words. These findings do not necessarily suggest that infants are born with knowledge of important grammatical categories. They do, however, show that infants’ perceptual biases facilitate dividing words into two fundamental categories. Subsequently, as infants approach the age of learning word meanings, they selectively focus on the louder and generally more salient content word category.

When they first begin to speak, infants typically omit function morphemes (e.g., *the*, *-ed*, *-s*), raising the question of whether infants simply do not perceive them or whether there are other reasons for the omissions, such as constraints on speech production (e.g., Gerken & McIntosh, 1993). Several lines of evidence suggest that infants do in fact perceive these weak items. For example, 11-month-olds, but not 10-month-olds, show a different pattern of scalp-recorded ERPs to stories that contained either correct or modified English function morphemes (Shafer, Shucard, Shucard, & Gerken, 1998). Similar results using the head-turn preference procedure with German infants suggest that 7- to 9-month-olds, but not 6-month-olds, can recognize previously familiarized closed class items (Höhle & Weisenborn, 2003). By 11 months of age, infants prefer nonsense words preceded by a familiar high frequency function word over that same nonsense word preceded by a mispronunciation of the function word (e.g., English infants listen longer to “the brink” over “ke brink”), and do so even for low frequency function words such as “its” or “her” by 13 months of age (Shi, Werker, & Cutler, 2003). Indeed, by 11 months of age it

80 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

appears that familiar high frequency function words such as “the” facilitate segmentation and learning of new words. Infants show better recognition of nonce word forms if they are first presented in a phrase with a familiar closed class item such as “the” (Shi, Werker, Cutler, & Cruickshank, 2004).

The importance of familiar function words becomes more pronounced when children are at the peak of learning new words, and then shows an apparent decline. In a preferential looking task (side by side pictures) infants of 18 months were most accurate when the labels for the objects were preceded by a correctly pronounced function word and least accurate when the function word was mispronounced. Their performance was intermediate for labels with missing function words. Infants of 24 months showed a similar, but less pronounced pattern, and by 36 months of age children were able to ignore the function word information (Zangl & Fernald, 2003). The role of function words extends beyond segmentation and identification. By 2 years of age infants are able to use function words as cues to new versus old information in a sentence context (Shady & Gerken, 1999).

Speech also carries paralinguistic (sometimes called “indexical”) information—cues that convey emotion, speaker identity, and emphasis—to which infants are sensitive. They show a preference for infant-directed over adult-directed speech (Cooper & Aslin, 1994; Fernald, 1984). They discriminate individual voices (DeCasper & Prescott, 1984; Floccia, Nazzi, & Bertoni, 2000) and show a robust preference for their mother's voice (DeCasper & Fifer, 1980) from birth, indicating prenatal learning. Indexical cues may also aid in the perception of specifically linguistic information. Phonetic discrimination, for example, is facilitated when the contrasting syllable is produced at the pitch peak in motherese, as shown by Karzon (1985) in infants' discrimination of /marana/ versus /malana/. This may be in part because the distinctiveness of segments may be clarified in motherese, as shown in the acoustic exaggeration of voicing (Ratner & Luberoff, 1984) and of the vowel space (Kuhl et al., 1997; Ratner, 1984) in infant-directed speech. Indeed, research shows that maternal clarification of the vowel space is correlated with superior speech discrimination in 6- to 12-month-old infants (Liu, Kuhl, & Tsao, 2003). Of interest, although speech directed to pets has many of the characteristics of infant-directed speech, the exaggeration of acoustic cues distinguishing vowels is evident only in speech directed to human infants (Burnham, Kitamura, &

Vollmer-Conna, 2003). This fact raises the possibility that the interplay between paralinguistic and linguistic factors may be part of, and exclusive to, within-species communication.

IMPLICIT DISCOVERY OF CUES IN THE INPUT: A DRIVE TO MAKE SENSE OF THE ENVIRONMENT

During the second half of the first year, an explosion occurs in infants' knowledge of detailed aspects of the sound structure of their native language(s), as noted in the earlier discussion of age related changes in phonetic perception. In this section, we document developmental changes in perception of other properties of language. The following section addresses the mechanisms that may be responsible for these learning trajectories.

Stress and Phonotactic Cues

Languages differ greatly in their internal prosodic regularities. By adulthood, speakers use these regularities to generate predictions about possible word structures. For example, English-speaking adults expect words to be trochaic—to begin with stressed syllables—mirroring the distribution of stress in their native language (e.g., Cutler & Carter, 1987; Cutler & Norris, 1998). This “trochaic bias” emerges early in the process of language acquisition, well before infants are producing words. For example, 9-month-olds prefer to listen to words that exemplify their native language's stress pattern (Jusczyk, Cutler, & Redanz, 1993), and are even sensitive to heavy versus light syllables (e.g., syllables with a long vowel and/or final consonant versus syllables with only a short vowel and no final consonant; Turk, Jusczyk, & Gerken, 1995). These results cannot be explained by recourse to inherent preferences for particular stress patterns, because 6-month-olds fail to show native-language stress preferences. This is an example of a potent learning process; somehow, English-learning infants must have discovered a probabilistic prosodic regularity in the input.

Infants' sensitivities are not confined to syllable-level patterns. By 9 months of age, infants have learned a great deal about the probabilistic phonotactic patterns of their native language, that is, the rates at which certain phoneme sequences occur in particular orders in particular positions in syllables and words. For example, the

sequence /ds/ can end, but cannot begin, syllables in English. Phonotactics are not a simple function of pronounceability; sequences that are legal in some languages are illegal in others. Effects of phonotactic structure are observed in studies of adult word recognition (e.g., Vitevitch & Luce, 1999; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). Similarly, children are affected by phonotactic probabilities when learning novel object names (e.g., Storkel, 2001) and nonword repetition tasks (Coady & Aslin, in press). By 9 months of age, infants prefer to listen to phonotactically legal sequences, whereas 6-month-olds do not (Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), with frequent phonotactic structures preferred over infrequent structures (Jusczyk, Luce, & Charles-Luce, 1994). Interestingly, infants in bilingual environments exhibit similar knowledge of phonotactic structure (Sebastián-Gallés & Bosch, 2002). Differences were obtained as a function of language dominance: infants were most sensitive to phonotactic patterns in their to-be-dominant language, suggesting that infants may be limited in the number of phonotactic systems they can acquire in parallel.

Some types of subsyllabic regularities appear to be more salient to infants than others. For example, Jusczyk, Goodman, et al. (1999) demonstrated that 9-month-olds were sensitive to sound patterns that recurred at the beginnings of words, but not the ends of words (for related findings, see Vihman, Nakai, DePaolis, & Hallé, 2004). These results suggest that certain parts of words may be privileged relative to others in infants' early speech representations, such that onsets may contain more detail than codas. Such findings are particularly interesting in light of the conventional wisdom that infants are highly attuned to rhyming, as well as data suggesting that the ends of words may be privileged in young children's lexical representations (Echols & Newport, 1992; Slobin, 1973).

Higher-Level Units

Infants' representations of the sound structure of their language also encompass larger prosodic patterns, spanning multiple words. Beginning in the 1980s, researchers have been interested in how such prosodic patterns might provide cues to infants to allow them to break into the syntax of their native language. Such prosodic bootstrapping accounts, beginning with Gleitman and Wanner's (1982) proposal for the use of weak

syllable function words as cues to grammar, have generally supported the claim that infants are attuned to the kinds of prosodic variables that are correlated with syntactic structure (for review, see Morgan & Demuth, 1996). One such prosodic cue is changes in pitch and duration at the ends of clauses in infant-directed speech (Fisher & Tokura, 1996; Jusczyk et al., 1992). In a classic study, Hirsh-Pasek et al. (1987) found that 7-month-olds listened longer to speech samples with pauses inserted at clause boundaries than sentences with pauses inserted clause-medially, suggesting that infants detected the disruptions in the latter case; similar results emerge for musical stimuli as well, suggesting that detection of prosodic markers serving as unit boundaries is not limited to language learners (Jusczyk & Krumhansl, 1993; Krumhansl & Jusczyk, 1990). More recent evidence suggests a similar process for phrase units, at least under some circumstances (Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003), despite less clear prosodic markers of phrases (Fisher & Tokura, 1996). Indeed, even newborns have been shown to be sensitive to cues correlated with prosodic boundaries (Christophe, Mehler, & Sebastián-Gallés, 2001).

Of course, this evidence doesn't demonstrate that infants "know" that these prosodic cues point to syntactic boundaries. And the prosodic phrases to which infants are sensitive correlate only imperfectly with syntactic boundaries (Gerken, Jusczyk, & Mandel, 1994). Nonetheless, this sensitivity is a prerequisite for the use of prosodic cues to discover syntactic structure. There is evidence from adult studies using artificial languages that such grouping cues do assist learners in breaking into syntax (e.g., Morgan, Meier, & Newport, 1987; Morgan & Newport, 1981). Other evidence suggests that prosodic structure helps infants as young as 2 months of age to organize and group word sequences in memory (e.g., Mandel, Jusczyk, & Kemler Nelson, 1994). However, the degree to which prosodic structure facilitates infants' discovery of syntactic structure remains unknown.

LEARNING MECHANISMS

It is of paramount importance to understand *how* the myriad information in the linguistic environment becomes part of the infant's native language knowledge base. Until the advent of new testing techniques, this question was only addressable via analyses of production

82 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

from older children, or via logical arguments regarding the structure of the problems facing the child and the possible solutions that might be part and parcel of the child's linguistic endowment (e.g., Pinker, 1984, 1989). Research with computational models also suggested possible ways to structure a learning system that might be compatible with some of the facts of language acquisition (Elman, 1990; Rumelhart & McClelland, 1986).

In the past decade, researchers have developed experimental methods that help to identify potential learning processes. Saffran, Aslin, and Newport (1996) used one such task to determine whether infants could track statistical properties of speech. Eight-month-old infants listened to a 2 minute continuous sequence of syllables containing multisyllabic words: for example, *golabupabikututibubabupugolabu. . .*. They subsequently tested infants' ability to discriminate the words from this "language" from syllable sequences spanning word boundaries. Infants' success at this task, as evidenced by different listening times to the two types of sequences, indicated that they were able to detect and use the statistical properties of the speech stream.

Similar methods can be used to test specific hypotheses about the types of learning mechanisms used by infants. For example, infants could have succeeded in the preceding task using two quite different types of statistics: the probabilities of co-occurrence of syllables (e.g., the transitional probability of "la" given "go"), or a simpler computation, frequencies of occurrence (test words occurred more often in the input than the other test sequences). Aslin, Saffran, and Newport (1998) teased apart these two possibilities, demonstrating that infants succeeded at this task even when the test items were matched for frequency of occurrence in the input. A recent computational analysis of infant directed speech confirms that probabilities of syllable co-occurrence predict word boundaries better than frequencies of syllable co-occurrences (Swingley, 2005).

We must still track frequencies to discover probabilities. And these frequencies are quite salient to infants in some linguistic domains. For example, infants represent the frequencies of phonotactic patterns in their native language (Jusczyk et al., 1994; Mattys, Jusczyk, Luce, & Morgan, 1999). Moreover, infants learn about frequent properties of the input before they learn about infrequent ones. Anderson, Morgan, and White (2003) showed that English infants show a decline at a younger age in their perception of the non-English retroflex-dental /da/-/Da/ distinction than the non-English velar-uvu-

lar /k/-/q/ distinction, presumably because "d's" are more frequent in the input than are "k's," giving infants a better opportunity to learn the native language category structure. Similarly, Shi et al. (2004) found that infants recognize and utilize high frequency function words earlier than infrequent ones.

Infant learning mechanisms are also sensitive to the statistical *distribution* of elements in the input. Maye, Werker, and Gerken (2002) presented 6- and 8-month-old infants with stimuli that simulated one of two types of languages. Materials in the unimodal condition collapsed a continuum of phonemes into a single category, such as English does with the two Hindi /d/ sounds, by presenting more instances of stimuli from the center of the continuum. Materials in the bimodal condition divided the continuum into two categories, such as Hindi does with the dental versus retroflex /d/, by including more instances of stimuli drawn from closer to the two ends of the continuum. The results suggested that infants were extremely sensitive to the distributions of the exemplars presented during exposure, with different test discrimination exhibited as a function of presentation of the unimodal versus bimodal materials. Distributional statistics can affect category structure, raising the possibility that sensitivity to the distributional information in the native language may contribute to the establishment of native language phonetic categories in the 1st year of life.

Infants also appear to be sensitive to nonstatistical regularities in the input. Marcus, Vijayan, Rao, and Vishton (1999) exposed infants to 3-syllable sentences following a particular pattern (e.g., *ga ti ga, li fa li*). Infants were then tested on novel sentences that either exemplified or violated the exposure pattern (e.g., *wo fe wo* versus *wo fe fe*). Successful discrimination suggested that the infants acquired abstract information, reflecting knowledge beyond just the specific syllable patterns observed in the input. Marcus et al. (1999) interpreted their results as evidence for a rule-based learning mechanism that detected algebraic rules (operating over variables). This claim has been controversial, as others have suggested that infants could have performed this task without rule-like representations (e.g., Altmann & Dienes, 1999; Christiansen & Curtin, 1999; Seidenberg & Elman, 1999). Investigators studying adults have similarly argued that the evidence supports a distinction between rule-based and statistical knowledge (Pena et al., 2003). However, it remains difficult to clearly distinguish between the two types of learning systems empir-

ically (for discussion, see Seidenberg, MacDonald, & Saffran, 2003).

Units for Computations

In order to specify the processes that go into the operation of any learning mechanism, it is necessary not just to note the structure of the learning mechanism (e.g., what computations are performed?) but also to determine the primitives over which those computations are performed. Consider the simplest possible mechanism, a frequency counter that tracks how often some event occurs. Depending on the event in question, the output of the learning mechanism could be vastly different. For example, if the mechanism is applied to a flock of birds, does it compute the total number of birds, or the number of birds' feet, or the number of swallows versus doves? Each of these primitives, serving as input to the learning process, renders a different answer.

The issue of primitives has been prominent in the study of speech representations. Artificial speech recognition systems intended to simulate early language development have largely focused on the phoneme as the relevant unit for modeling (e.g., Brent & Cartwright, 1996; Christiansen, Allen, & Seidenberg, 1998; Jusczyk, 1997). Although some research supports this unit as important in infant speech perception, other work suggests that this idealization may be a mismatch to infants' capabilities.

Shortly after the field of infant speech perception emerged, researchers began to focus on what the unit of representation might be. One long-standing controversy concerns whether syllables or phonemes (or both) are psychologically real to infants. Studies using discrimination tasks in which either the syllable changes, or a phonetic feature in the segment changes, provided convincing evidence that both syllable-level and segment (phoneme) level features are accessed and used by 2- to 4-month-old infants (Eimas & Miller, 1981; Miller & Eimas, 1979). However, studies using similarity assessments (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988; Jusczyk & Derrah, 1987) have yielded a different pattern of results, suggesting that young infants are sensitive to changes in the number of syllables in a word, but not the number of phonemes (Bijeljac-Babic et al., 1993). Despite the contradictions in the studies with younger infants, older infants include some subsyllabic structures in their representations (Jusczyk, Goodman, et al., 1999). Moreover, with development

and/or literacy, adults represent both syllables and phonemes (Nygaard & Pisoni, 1995), and recent adult studies suggest that segmental representations may serve as the primitives for at least some kinds of language learning tasks (Newport & Aslin, 2004). Of additional interest, there may be differences across languages in which unit adults use for word segmentation, with French adults showing a pronounced syllable bias and English adults showing sensitivity to segmental information as well (Cutler, Mehler, Norris, & Segui, 1983, 1986).

One interpretation of this body of work is that the syllable is privileged as a unit of representation (Bertoncini & Mehler, 1981), and may be used as the unit in computations across linguistic input. Another interpretation is that both syllables and phonemes are privileged, but for different types of tasks: the syllable is the primary unit for counting, but segmental detail plays a role in segmentation, at least in stress-timed languages (see Werker & Curtin, 2005). Further research will help distinguish between these possibilities.

BUILDING FROM THE INPUT DURING THE 1ST YEAR

The foregoing review provides some clues regarding the types of linguistic information infants acquire during the 1st year, as well as potential learning mechanisms that subserve this acquisition process. We can now turn to the burgeoning literature that puts these two pieces together: studies concerning the acquisition of particular linguistic structures. Some of these studies teach infants new information during a laboratory exposure session. Other studies ask how infants use what they've previously learned about their native language to discover structure in novel input.

Many of these studies use artificial nonsense languages, a methodology taken from the adult literature, in which specific cues can be isolated in a way that is impossible in natural speech (e.g., Gómez & Gerken, 2000; Morgan et al., 1987). Such languages are particularly useful in infant studies because they permit the development of brief exposure materials, fitting the task demands of infants with limited attention spans. On the other hand, artificial materials sacrifice ecological validity. An issue currently confronting researchers is the need to demonstrate that the learning abilities uncovered using artificial methods are the same as those infants use when acquiring their native language. While

84 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

ecological validity is always an issue in laboratory learning studies, it is particularly salient given exposure regimens that are so clearly unlike natural language input.

Learning Phonology and Phonotactics

Phonotactic knowledge is a prime candidate for learning studies, because it is so clearly tied to the structure of particular languages. Moreover, phonotactics is somewhat different from many of the other features of language that infants acquire. Phonotactic patterns are both general (they apply across the whole language, and are not specific to known words) and specific (they consist of segmental patterns, unlike the syllabic and prosodic patterns that often appear to be the focus of infants' attention). Studies of phonotactics in adults suggest a learning process in which mere exposure to phonotactic regularities influences expectations about possible word form regularities (Dell, Reed, Adams, & Meyer, 2000; Onishi, Chambers, & Fisher, 2002).

Chambers, Onishi, and Fisher (2003) extended their adult studies to include 16.5-month-old infants. The materials consisted of nonsense word sequences in which consonant positions were restricted; for example, /b/ could occur word-initially but not word-finally. Following exposure, infants listened longer to syllables that were phonotactically legal than those that violated the exposure patterns. Impressively, infants were also able to learn 2nd order phonotactic regularities, in which the presence of one element was conditioned on the presence of the other (e.g., that /k/ begins syllables if and only if the subsequent vowel is /ae/).

To determine whether certain phonotactic regularities are harder to learn than others, Saffran and Thiessen (2003) exposed infants to two different types of phonotactic patterns. One was consistent with the types of patterns found cross-linguistically, while the other was unlike natural language structure. Infants rapidly learned regularities of the first type, which involved generalizations across sets of acoustic/linguistic features (such as voicing, the feature that clumps /p/, /t/, and /k/ into a separate category from /b/, /d/, and /g/). However, infants failed to learn regularities that disregard such linguistic features (such as the grouping of /p/, /d/, and /k/ versus the grouping of /b/, /t/, and /g/), which are unlike natural language patterns. These results suggest a possible explanation for *why* languages show the types of patterning that they do. Sound structures that are hard for infants to learn may be less likely

to recur cross-linguistically. More generally, studies that uncover infants' failures may turn out to be as illuminating as those that display infants' considerable strengths at learning, by highlighting constraints on infant learning mechanisms.

The acquisition of phonological knowledge is also of great interest, given the rapidity with which infants acquire such knowledge in their native language. This learning process requires infants to integrate different types of information, an ability that likely emerges between 6 and 9 months of age (e.g., Morgan & Saffran, 1995). For example, consider the trochaic bias—the expectation that (at least for English-learning infants) words begin with stressed syllables. To learn this pattern, you must know something about the relationship between stressed syllables and their positions within words. If you have yet to discover any word boundaries, it would be impossible to know that stressed syllables fall in predictable places in words. Thus, to acquire a trochaic bias, infants must learn correspondences between stress and word position. This is only possible if infants first know some trochaic words, which may explain the lack of such a bias in 6-month-old infants. Indeed, 6½-month-old infants can be taught a rhythmic bias by briefly exposing them to word lists exemplifying the bias, and 9-month-olds' biases can be similarly altered (Thiessen & Saffran, 2004).

Other studies investigate how infants acquire more abstract phonological knowledge concerning the stress assignment patterns of their native language. Gerken (2004) presented infants with artificial language stimuli designed to exhibit particular patterns of metrical phonology—the structural principles for stress assignment in multisyllabic words. Following a brief exposure to a word list in which certain stress assignments were exemplified, infants were tested to determine whether they had inferred stress pattern structures that had not actually occurred in the input. The results suggest that 9-month-olds generalize to new words using abstract knowledge of possible stress patterns, opening the door to additional studies probing the extent to which infants are able to learn the types of abstract phonological structures that typify human languages.

Word Segmentation

The problem of how infants discover words in fluent speech, which lacks consistent physical cues to word boundaries (Cole & Jakimik, 1980) has played a promi-

ment role in studies of early language learning. While interest in this problem is a relatively recent development in the field of language acquisition, there are some notable exceptions. Roger Brown began his classic 1973 volume on language acquisition by describing his own problems with word segmentation while taking a Berlitz course in Japanese; later he described an early model of distributional learning in word segmentation by Olivier (1968). Gleitman and Wanner (1982) also treated the problem seriously, hypothesizing that stressed syllables may mark words for young learners. These two early discussions of the segmentation problem, invoking distributional and prosodic cues, were prescient, as these two sources of information are currently at the forefront of theories regarding infant word segmentation.

In a seminal study, Jusczyk and Aslin (1995) used the head-turn preference procedure to determine when infants begin to segment words. They first presented 7½-month-old infants with a word segmentation problem: sentences in fluent speech containing a particular target word (e.g., “*Mommy’s cup is on the table. Do you see the cup over there?*”). Following this familiarization period, infants were tested on the target words (e.g., “*cup*”) versus novel words (e.g., “*bike*”). Each item was played for as long as the infant maintained a head-turn in the direction of a speaker from which the word was played. Jusczyk and Aslin (1995) found a significant difference in listening times between the familiar and novel words, suggesting that the 7½-month-olds discovered the target words in fluent speech. Six-month-olds, however, failed to show any significant differences between the familiar and novel test items, suggesting either that the ability to segment word from fluent speech develops sometime between 6 and 7½ months of age, or that younger infants require additional exposure and/or cues to successfully perform the task. Support for the latter view comes from a study by Thiessen and Saffran (2003), in which 6½- to 7-month-olds successfully performed a word segmentation task in which they received more familiarization with the target words, and a study by Bortfeld, Morgan, Golinkoff, and Rathbun (2005) demonstrating word segmentation by 6-month-olds using additional cues.

How do infants solve such a complicated task? A growing body of evidence suggests that infants are attuned to a number of cues correlated with word boundaries. One such source of information was initially suggested in the linguistics literature in the mid-twentieth century (e.g., Harris, 1955), reflecting the observation that words consist of predictable sequences of

sounds. To see this statistical structure, consider the following example: because the syllable *pre* precedes a small set of syllables in English, the probability that *pre* is followed by *ty* is quite high. However, because the syllable *ty* occurs word-finally, it can be followed by any syllable that can begin an English word. Thus, the probability that *ty* is followed by *ba*, as in *pretty baby*, is extremely low. Indeed, infants are sensitive to such probabilistic cues, and use them for word segmentation (e.g., Aslin, Saffran, & Newport, 1998; Goodsitt, Morgan, & Kuhl, 1993; Saffran et al., 1996).

Several lines of research have converged to suggest that particular languages contain prosodic cues that facilitate word segmentation. For example, English-learning 7½-month-olds can make use of their knowledge that bisyllabic words tend to be trochaic to successfully segment strong-weak words (those stressed on their first syllable) such as “KINGdom” from fluent speech, while failing to segment weak-strong words (those stressed on their second syllable) like “guiTAR” (Jusczyk, Houston, & Newsome, 1999). In the latter case, infants treat the stressed syllable “TAR” as a word. Interestingly, they will combine TAR with a subsequent weak syllable if it appears consistently, suggesting the integration of stress-based and statistically based strategies. Thus, infants use their expectations about word structure to assist in segmentation (for related results, see also Curtin, Mintz, & Christiansen, 2005; Houston, Jusczyk, Kuijpers, Coolen, & Cutler 2000; Houston, Santelmann, & Jusczyk, 2004; Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, in press).

Younger infants are unable to take advantage of stress-based segmentation cues, demonstrating that this knowledge must be learned (e.g., Echols, Crowhurst, & Childers, 1997; Jusczyk, Houston, et al., 1999; Thiessen & Saffran, 2003). Further evidence for a learning account comes from research on languages that incorporate different stress patterns, such as French (Polka, Sundara, & Blue, 2002). Artificial language studies also indicate that stress-based segmentation strategies are learnable (Thiessen & Saffran, 2004). Moreover, infants must learn not to overly focus on stress, which, like all cues to word boundaries in isolation, is fallible; only by 10½ months do infants successfully segment weak-strong words (Jusczyk, Houston, et al., 1999).

The fact that infants can use the distribution of stress cues as a cue to word boundaries raises a “chicken-and-egg” problem. If stress is a critical cue to word boundaries, how can infants have discovered the utility of this

86 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

cue prior to knowing words? One must know something about the words of one's native language to discover the correlation between stress position and word boundaries. One possibility is that infants learn the predominant stress pattern of words in their native language by hearing words spoken in isolation (e.g., Jusczyk, Houston, et al., 1999). This seems intuitively plausible, particularly given analyses of infant-directed speech suggesting that a nontrivial proportion of utterances consist of single words (e.g., Brent & Siskind, 2001). On this view, infants might learn words like "kitty" and "mommy" by hearing them spoken in isolation, and then use that nascent corpus to discover the stress patterns characteristic of their native language. However, a recent analysis suggests that this explanation is unlikely to be correct (Swingley, 2005). Only 14% of the bisyllabic utterances in English spoken to infants are trochaic; most bisyllables conform to a strong-strong pattern. Thus, infants must have some other means of discovering the predominant lexical stress pattern of their native language.

One possibility is that early access to sequential statistical segmentation cues provides infants with the beginnings of the corpus they need to subsequently discover prosodic regularities. Swingley's (2005) computational analysis suggests that words that are discoverable via statistical cues render the correct prosodic template, unlike words heard in isolation. A study by Thiessen and Saffran (2003) suggests a trajectory of cue usage over development consistent with this view. When 9-month-olds are confronted with continuous speech in which stress and statistical cues conflict, they follow the stress cues, as previously demonstrated by Johnson and Jusczyk (2001). However, 6- to 7-month-olds exhibit the opposite strategy, relying on statistical cues rather than stress cues, presumably because they do not yet know their native language's stress pattern. The stress strategy, then, is presumably bootstrapped from the regularities in the initial corpus acquired via sequential statistical cues (Thiessen & Saffran, 2004).

Other important cues to word boundaries become available to infants beginning around 9 or 10 months of age. For example, infants are able to use the distributions of allophones—the subtle differences in phonemes that are a function of the context in which the phoneme occurs—as word boundary cues. Certain sounds only occur in certain positions in words; the /t/ that begins English words differs from the /t/ that occurs word-medially or word-finally (Church, 1987). Young infants are sensitive to allophonic cues which might signal word boundaries, shown, for example, by their ability to dis-

criminate the bisyllable /mati/ when the /ma/ and the /ti/ are pulled from either a single word or from the final syllable in one word versus the first syllable in the next (Christophe, Dupoux, Bertoncini, & Mehler, 1994). By 9 months of age, infants can detect word boundaries in contrasts such as *nitrates* versus *night rates*, which consist of the same sequence of phonemes but different allophones, suggesting the availability of allophonic cues for segmentation (Jusczyk, Hohne, & Bauman, 1999; Mattys & Jusczyk, 2001). Moreover, at this same age infants' phonetic categories reflect sensitivity to position-specific allophonic variants (Pegg & Werker, 1997). These findings raise the same sort of "chicken and egg" problem as the stress findings—one must first know something about words to discover cues correlated with internal word structures. It is thus likely not an accident that this ability emerges at roughly the same time for different types of cues. By 9 months of age, infants have likely segmented enough words using statistical cues and other types of information to have developed a sufficiently large corpus to discover these word-internal cues.

Phonotactic cues are also correlated with word boundaries (e.g., Brent & Cartwright, 1996; Cairns, Shillcock, Chater, & Levy, 1997; Vitevitch & Luce, 1998). For example, Mattys et al. (1999) demonstrated that infants use the likelihood that particular consonant clusters occur within or between words in their native language as a segmentation cue. Infants' ability to segment sequences such as *nongkuth* versus *nomkuth* was examined. Critically, while the consonant clusters in the middle of each sequence are equally likely in English, the former is more likely to occur within words (/ngk/) while the latter is more likely to span a word boundary (/mk/). Nine-month-olds used this subtle distinction as a segmentation cue, inferring word boundaries in the middle of *nomkuth* but not *nongkuth*. A segmentation strategy based on phonotactics requires the infant to already know enough words for these regularities to become apparent. Related segmentation cues may require no prior lexical experience. For example, 12-month-old infants follow the "Possible Word Constraint": they generate segmentations that only create possible words, while avoiding stranding sequences that are not possible words, such as sequences consisting of a single consonant (Johnson, Jusczyk, Cutler, & Norris, 2003). This constraint may help infants to segment speech appropriately and to avoid errors without requiring a lexicon from which to induce the constraint.

It should be clear at this point in the discussion that no single cue underlies word segmentation. This conclu-

sion is evident both from the empirical literature demonstrating that infants are sensitive to myriad cues and from the fact that each cue, in isolation, only solves part of the problem for infants. Studies using multiple cues have largely asked how infants weight conflicting cues. For example, 6- to 7-month-olds prioritize statistics over stress, while 9-month-olds prioritize stress over statistics (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003). Interestingly, Mattys et al. (1999) found that 9-month-olds also prioritize stress over phonotactic cues, supporting the hypothesis that, while imperfect, stress cues are relatively easy to detect and use (Thiessen & Saffran, 2003); Mattys et al. (1999) suggest that “prosody is an initial cue yielding a coarse first pass at word boundaries that is subsequently supplemented with additional cues such as phonotactic and allophonic constraints” (p. 482). However, it remains unknown how such cues are combined in infants’ emerging segmentation strategies (see Morgan & Saffran, 1995, for an example of a study looking at additive effects of cue combinations).

One avenue of research that has effectively explored the use of cue combinations for the discovery of word boundaries is the computational literature (for an extensive review, see Batchelder, 1997). For example, Christiansen et al. (1998), building on the work of Aslin, Woodward, LaMendola, and Bever (1996), examined the efficacy of phonotactic cues that predict ends of utterances as a cue to word boundaries in a corpus of child-directed speech. While this cue worked only moderately well in isolation, inclusion of lexical stress cues markedly improved the performance of the network. A different approach to this problem was pursued by Curtin et al. (2005), who found that including stress information in a corpus enhanced performance by allowing the network to represent stressed and unstressed variants of the same syllable as distinct. One of the messages provided by the computational literature is that more cues are probably better than fewer cues, despite the paradoxical fact that this makes the input more complex. Awaiting future research is the determination of exactly which cues infants attend to, and whether these cues are weighted in the manner predicted by the computational models.

Throughout this section, we have been discussing word segmentation as though it is clear that infants are discovering words in the input, and subsequently representing these sound sequences as units, available for later mapping to meaning. However, it is certainly possible that infants are engaged in a simpler process. Return-

ing to the original Jusczyk and Aslin (1995) study, we earlier described the results as evidence that infants had segmented the word “cup” from the fluent speech. It is equally possible, though, that infants’ test performance—discriminating “cup” from “bike”—rests on simply recognizing that the former set of sounds is more familiar than the latter. Doing so would not necessitate segmentation per se; instead, infants would be responding based on the familiarity of the sounds, without having represented “cup” as a distinct lexical representation. Indeed, one early study suggested that in segmentation tasks, infants pull out metrical feet (a rhythmical unit) rather than actual words (Myers et al., 1996). It is thus of great interest to ask what the output of word segmentation actually is. Saffran (2001) addressed this issue with respect to the statistical learning results. When infants respond to *golabu* during testing, after exposure to *golabupabikututipugolabu . . .*, are they treating *golabu* as a word, or as a familiar sound sequence? Based on results from a task in which infants are tested on words like *golabu* embedded in English sentences after exposure, Saffran (2001) suggested that infants treat these nonsensical patterns as primitive English words (i.e., whatever a word is to an 8-month-old, in the absence of mapping to meaning).

Recent studies by Curtin et al. (2005) with 7-month-old infants further suggest that stress cues are represented in these newly segmented proto-lexical representations. Curtin et al. used analyses of child-directed speech to argue that infant learners would be more successful if they represent stressed and unstressed syllables differently during word segmentation. In particular, the analyses suggest that incorporation of stress into infants’ representational landscape would result in better distribution-based word segmentation, as well as an advantage for stress-initial syllable sequences. Results of a behavioral study corroborated these analyses. In particular, if items in the test phase were placed in a sentence context and the target was either an exact match (*BEdoka*) as opposed to a sequence with the same segments but a different stress pattern (*beDOka*) or other type of nonmatching control sequences, infants demonstrated an overwhelming preference for the exact match. These results suggest that stress information in the ambient language not only shapes how statistics are calculated over the speech input, but that it is also encoded in the representations of parsed speech sequences.

Once some sequences have been segmented from the speech stream to become new lexical entries, can these

88 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

words assist in segmentation of subsequent fluent speech, helping infants to discover other adjacent words (e.g., Brent & Cartwright, 1996; Dahan & Brent, 1999)? A recent study by Bortfeld et al. (2005) provides evidence that 6-month-old infants can use known words to segment new words from fluent speech. Infants heard continuous speech in which the word to be segmented appeared adjacent to the infant's own name, which infants recognize early in the 1st year (Mandel, Jusczyk, & Pisoni, 1995). The familiar name served as a strong segmentation cue, providing the first positive evidence for word segmentation in infants as young as 6 months. By demonstrating that infants' prior knowledge alters the manner in which they process new input, the Bortfeld et al. (2005) results suggest a promising new tact for studies of word segmentation and infant learning more generally.

Beginnings of Word Recognition

Once infants have segmented words into discrete units, they are ready to begin recognizing familiar words, matching internal representations of words to their instantiation in subsequent input. This is no simple matter, because words are not static invariant patterns. The sounds of any given word are shaped by properties of the speaker (such as speaker's voice, sex, speaking rate, and affect) and by the context in which the words are produced (such as coarticulation effects).

What words might one expect infants to first recognize? Mandel et al. (1995) hypothesized that infants' names might be particularly salient. They occur frequently, are often presented in isolation, and likely carry affective prosody that attracts infants' attention. Using the preferential listening procedure, 4½-month-olds heard either their own name or an unfamiliar name. Infants preferred to listen to their own names, suggesting that they matched internal representations of these familiar sounds to the input played during the experiment. This does not mean that these infants knew the meanings of these sounds. However, by 6 months of age, infants can recognize highly familiar words based on their meanings. When presented with side-by-side video displays of their mother and father, infants look longer to the display that matches auditory presentations of "mommy" versus "daddy" (Tincoff & Jusczyk, 1999).

These results suggest that infants' developing lexical representations are not fleeting, but are built up incrementally and maintained over time, despite the variabil-

ity in the input. To explicitly investigate the time course of memory for new lexical representations, Jusczyk and Hohne (1997) exposed 8-month-olds to stories containing particular vocabulary items. After 10 days of exposure to the stories, a 2-week retention interval was introduced, during which infants did not hear the stories. Infants were then tested on their recognition of words from the stories versus similar words that had not occurred in the stories. Despite the 2-week retention interval, during which infants heard a vast array of potentially interfering speech, the infants listened longer to the words from the previously familiarized stories, suggesting that auditory representations that were garnered weeks before were sufficiently robust to support later word recognition.

How detailed are these early representations? Infants appear not to confuse similar sounding words, at least under certain circumstances. Infants in Jusczyk and Aslin's (1995) experiments did not incorrectly treat "zeet" as familiar after being exposed to "feet," suggesting that these early representations are fairly specific. Similarly, after repeated exposures to a word and object, 8-month-old infants show robust evidence of detecting a change to a new word that differs in only a single phonetic feature (Stager & Werker, 1997).

Early word representations also appear to include a level of acoustic detail that corresponds to the positions of syllables relative to structural boundaries in sentences, such as phonological phrases. Acoustic cues corresponding to phonological phrase boundaries are detected even by newborn infants (e.g., Christophe et al., 1994, 2001). By 13 months of age, infants can use this distinction in word recognition (Christophe, Gout, Peperkamp, & Morgan, 2003). For example, infants trained on the word "paper" were tested on sentences in which paper was either a word ("The college with the biggest *paper* forms is best") or in which paper spanned a phonological phrase boundary ("The butler with the highest *pay* performs the best"). Despite the fact that the syllable sequence was the same in both cases, with equivalent statistical and stress cues, the results suggested that the infants' representations included the subtle acoustic differences between "paper" and "pay per."

One source of information that may aid infants in speech processing is coarticulation. In order to produce speech as rapidly as we do, whenever we produce a segment, syllable, or word, we move the lips, tongue, and jaw in a way that maintains the positions required for that segment as well as for both the preceding and fol-

lowing consonants and vowels. For example, because of coarticulation, the phoneme /b/ is different in the word “beet” than in the word “boot.” Adults are sensitive to this coarticulatory information, but only under some listening conditions. For example, when words in a string are presegmented by the insertion of pauses, adult listeners show better recognition of those familiar syllables that maintain the same coarticulatory information as used during familiarization. However, when the pauses are omitted and adults must rely on only transitional information, their access to coarticulatory information is no longer evident (Curtin, Werker, & Ladhar, 2002). Seven-month-old infants are also sensitive to coarticulatory information, but under the opposite conditions as adults. When the syllables are presegmented by the insertion of pauses, infants’ recognition of familiar words is not enhanced by matching coarticulatory cues. However, in tasks that require the infant to segment syllables from a continuous stream of speech, matching coarticulatory information significantly improves performance (Curtin et al., 2002).

Infants also appear to represent indexical information that affects word recognition. For example, 7½-month-old infants readily recognize words previously heard produced by a speaker of the same sex, but show no evidence of word recognition when the target is produced by a speaker of the opposite sex (Houston & Jusczyk, 2000). It seems likely that infants’ representations include perceptual features of the speaker, such as components of pitch, that make cross-sex matching challenging. Similarly, 7½-month-olds represent the affective state of the speaker, showing word recognition only when the affective state of the familiarized words matched the targets (Singh, Morgan, & White, 2004). Related arguments are emerging in the field of infant music perception, where researchers are actively investigating the “grain” at which infants represent musical experiences in memory for subsequent recognition (e.g., Ilari & Polka, 2002; Palmer, Jungers, & Jusczyk, 2001; Saffran, Loman, & Robertson, 2001; Trainor, Wu, & Tsang, 2004).

Listening for Meaning

In contrast to the detailed, multiple levels of information available to prelinguistic infants in word recognition and segmentation tasks, infants who have begun to assemble a more sizeable lexicon seem to be more selective and more limited in which detail they use to recog-

nize words. For example, although 14-month-olds can learn to associate two different nonsense words with two different objects (Schafer & Plunkett, 1998; Werker, Cohen, Lloyd, Casasola, & Stager, 1998; Woodward, Markman, & Fitzsimmons, 1994), they fail at this same age if the two nonsense words are phonetically similar such as “bih” and “dih” (Stager & Werker, 1997) or “pin” and “din” (Pater, Stager, & Werker, 2004). Importantly, 14-month-old infants succeed in a virtually identical task when the word is paired with a visual display that is unlikely to evoke labeling (Stager & Werker, 1997). Moreover, when an easier variant of the task was used in which a single object was paired with a single word, the 14-month-olds still failed to notice the change to a phonetically similar word, whereas 8-month-olds succeeded in this same task. The failure to learn minimally contrastive words was shown to be short-lived. When tested in exactly the same task, 17- and 20-month-old infants succeeded at learning phonetically similar words (Werker, Fennell, Corcoran, & Stager, 2002), as did even infants of 14-months who had particularly sizeable vocabularies (Werker et al., 2002; see also Beckman & Edwards, 2000 for a discussion of the potential role of vocabulary size). An identical pattern of results was obtained using an ERP paradigm in which a higher amplitude deflection is seen to known versus unknown words (Mills et al., 2004).

Why might 14-month-old infants fail to distinguish phonetically similar words in a word-learning task when they could still discriminate these two words, and when both younger and slightly older (or even more advanced same-aged infants) succeed? Stager and Werker (1997) speculated that for the novice word learners, the computational demands of linking a word with an object are so great that the attentional resources are not available to utilize all the word-level detail that is perceived (see Kahneman, 1973, for the original postulation of attention as a limited resource). However, other interpretations of these findings were that they revealed evidence of a discontinuity between the representations used in phonetic versus phonological (or lexical) representations. Indeed, there is a long-standing tradition in child phonology that posits such a representational discontinuity (see Brown & Matthews, 1997; Rice & Avery, 1995; Shvachkin, 1948). Empirical work by Hallé and de Boysson-Bardies (1994) provided potential support for this discontinuity hypothesis. They used a word recognition task requiring infants to listen to lists of highly familiar versus unknown words, and found that although

90 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

infants of 7 and 11 months both showed a preference when the unknown words were phonetically dissimilar from the known words (Hallé & de Boysson-Bardies, 1994), if phonetically similar foils were used, only the 7-month-old infants succeeded.

A number of subsequent studies have now disconfirmed the discontinuity hypothesis. When tested in a simpler word recognition procedure wherein infants are shown two pictures and presented with a single word that either matches one of the objects or is a mispronunciation of the same name for that object (e.g., “baby” versus “vaby”), infants from 20 (Swingley & Aslin, 2000) down to 14-months (Swingley & Aslin, 2002) can detect the mispronunciation. Sometimes this detection is shown in longer looking to the correct object when the word is pronounced correctly (Swingley & Aslin, 2002), and sometimes it is evident in a shorter latency to look away from the mismatch (Swingley & Aslin, 2000), but it is consistently evident. This success is seen for well-known words in the associative task used by Stager and Werker (1997). If habituated to the word “ball” paired with the moving object “ball,” and the word “doll” paired with a visual display of a “doll,” 14-month-old infants detect a switch in the word object pairing in the test phase (Fennell & Werker, 2003). Thus, as suggested by both Stager and Werker (1997) and Swingley and Aslin (2000), it appears that an attentional resource limitation rather than a representational discontinuity accounts for the failure of 14-month-old infants under some circumstances.

An attentional resource limitation may not fully explain the preceding findings. It is still necessary to know why it is that phonetic detail is dropped. Is this the only detail that infants drop at 14 months, or is other detail also ignored? The word segmentation and recognition studies revealed that 7- to 9-month-old infants utilize many different kinds of information in the signal. This is evident, for example, in their failure to recognize words if there is a change in speaker gender (Houston & Jusczyk, 2000), or affect (Singh, Bortfeld, & Morgan, 2002). However, by 10½ months of age, infants successfully recognize words spoken by opposite-sex speakers (Houston & Jusczyk, 2000), and are able to ignore changes in affect and still show evidence of recognizing familiar words (Singh et al., 2002). One increasingly popular account, which is somewhat different from the traditional view in which indexical information is not part of lexical representations, is that a rich tapestry of

information—phonetic, indexical, coarticulatory—is included in the lexicon (e.g., Goldinger, 1992), but that not all of this information is used in every task situation (Werker & Curtin, 2005).

Evidence in support of this possibility is provided in recent work with 2.5- and 3-year-olds by Fisher, Church, and Chambers (2004). They demonstrated that children represent both abstract and detailed linguistic information pertaining to the specifics of pronunciation of familiar words. For example, their participants represented the distinction between a medial /t/ and a more /d/-like flap pronunciation of the same phoneme, despite the fact that both pronunciations are legal. Interestingly, the same pattern of results emerged in a related study using nonwords, suggesting that even new lexical representations—formed after just a few exposures to a word—are flexible, in that they are both abstract and specific (Fisher, Hunt, Chambers, & Church, 2001). These findings suggest that perceptual learning mechanisms used flexibly throughout life to adapt to new linguistic input may operate from the beginning of the word learning process (Fisher et al., 2004).

It may be that in the earliest stages of word learning, infants are less able to flexibly select which information to attend to. With the attentional resource demands of attaching meaning to words, infants at the cusp of word learning may be captured by that information which is most salient. To test this hypothesis, Curtin and Werker (cited in Werker & Curtin, 2005) recently tested the ability of 12-month-old infants to learn words that are similar in all respects except stress pattern. They found that these infants, a full 2 months younger than the infants who failed to learn phonetically similar words, could successfully learn to map words such as DObita versus doBIta (where capitals indicate stress) onto two different objects.

With these studies, the links between infant speech perception, word segmentation, word recognition, and word learning are being much more fully described. Moreover, the infant literature is beginning to interface much more richly with the large literature on adult lexical access. Infancy researchers are no longer restricted to asking questions like “what is the unit of representation”? Instead, the field is now poised to allow the asking of much more nuanced questions such as “what information is utilized, when, and why?”

The advent of new methodologies has allowed researchers to go beyond asking which words infants rec-

ognize to assess the time course of word recognition. Eye-tracking has become an important tool in assessing adults' lexical representations (e.g., Allopenna, Magnusson, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Adapting these methods to study infants, Fernald, Pinto, Swingley, Weinberg, and McRoberts (1998) assessed infants' speed and accuracy at word recognition over the course of the 2nd year. To do so, the experimenters measured infants' eye movements as they viewed computer-displayed pictures of familiar objects while listening to the names of these objects. Speed and reliability were correlated with age, suggesting that infants' lexical representations likely become more robust and the cognitive machinery underlying word recognition becomes more fluent during the 2nd year. Like adults, 24-month-olds do not need to hear an entire word to recognize it; instead, word recognition is incremental (Swingley, Pinto, & Fernald, 1999). For example, these infants rapidly distinguished *doggie* from *tree*, correctly fixating on the matching picture, but took 300 ms longer to distinguish *doggie* from *doll*, reflecting the increased phonetic overlap of the latter pair. Interestingly, infants recognize parts of words just as rapidly as whole words, supporting the view that infants, like adults, process words incrementally (Fernald, Swingley, & Pinto, 2001). This ability appears to be associated with infants' productive vocabularies, suggesting a link between lexical growth and the efficiency with which infants recognize words. These facts about how infants process words are consistent with corpus analyses suggesting that the words in infants' early vocabularies are sufficiently overlapping in phonological space to necessitate detailed lexical representations (e.g., Coady & Aslin, 2003).

Experience with particular words appears to enhance these nascent lexical representations. Church and Fisher (1998) observed long-term auditory priming in 2- to 3-year-olds very similar to that of adults, showing effects of experience with specific words on subsequent word identification and repetition. Similar effects emerged in a study with 18-month-olds using a preferential looking task, suggesting that just two repetitions of a word assisted infants in subsequently identifying the target word (Fisher et al., 2004). Similarly, neuroimaging tasks are broadening the range of questions that can be explored. For example, infants may fail to discriminate a nonnative contrast given the task demands of a behavioral task, yet still show evidence of a neurophysiologi-

cal response to the change, indicating that at some level in the brain the information is available (Rivera-Gaxiola et al., 2005).

Beginnings of Grammar

Since most infants do not begin combining words grammatically until the ripe old age of 18 to 24 months or beyond, is there any reason to suspect that the capacity to acquire grammatical structure is present earlier in life? Indeed, researchers have demonstrated early evidence for grammatical knowledge of the native language during infancy, as well as precocious abilities to learn new, simple, grammatical structures using artificial language methodologies (see Tomasello, Chapter 6, this *Handbook*, this volume, for a review of the literature on subsequent aspects of grammar learning). Comprehension studies suggest that infants have a sophisticated grasp of certain syntactic structures by the end of the 2nd year. For example, Naigles (1990) tested young 2-year-olds in a cross-modal matching task that required them to induce the meaning of a new verb. The infants heard either transitive structures, such as "The duck is kradding the bunny," or intransitive structures, such as "The duck and bunny are kradding." Infants looked longer at a video that matched the syntactic structure of the sentence they heard. These results suggest that infants can engage in what is known as "syntactic bootstrapping": using their prior knowledge of syntactic syntax (here, transitivity) to determine the meaning of *kradding*.

Infants' morphological knowledge is similarly advanced. For example, Santelmann and Jusczyk (1998) exposed infants to passages that contained either a grammatical English dependency between the auxiliary verb *is* and a main verb ending with *-ing*, or an ungrammatical combination of the modal auxiliary *can* and a main verb ending with *-ing*. Eighteen-month-olds, but not 15-month-olds, discriminated between the two types of passages. These results suggest that by the middle of the 2nd year, infants have learned how certain types of discontinuous grammatical dependencies operate in their native language.

A number of recent studies have employed artificial grammar methodologies to uncover the learning mechanisms underlying this process. When exposed to word sequences ordered by simple rules (e.g., Marcus et al., 1999) or finite state grammars (Gómez & Gerken, 1999), infants treat test items that violate those patterns

92 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

as novel, even if they are instantiated in new vocabulary. For example, Gómez and Gerken (1999) used the head-turn preference procedure to assess whether 12-month-olds could acquire a miniature artificial grammar. Infants discriminated new grammatical strings from ungrammatical strings after less than 2 minutes of training, with evidence that they acquired both specific information (e.g., legal beginnings and ends of sentences, and internal pair-wise combinations) and abstract information (e.g., grammatical structures produced using a new set of vocabulary). Ongoing research is probing the circumstances under which infants are more or less likely to generalize beyond the input given (e.g., Gómez, 2002; Gómez & Maye, 2005).

Saffran and Wilson (2003) extended this line of research to ask how infants might approach learning tasks consisting of multiple levels of information. Twelve-month-olds listened to a continuous speech stream in which the words were ordered via a finite-state grammar. The infants were thus presented concurrently with a word segmentation task and a syntax learning task. The results suggest that infants can first segment novel words and then discover syntactic regularities relating the new words—all within the same set of input. Studies of this type indicate that artificial learning situations can be scaled up to begin to represent some of the problems confronting learners faced with natural language input. For example, Gerken, Wilson, and Lewis (2005) performed a hybrid artificial/natural language learning study in which infants heard a small subset of Russian words marked with correct gender morphology. The results demonstrate that certain types of patterns that occur in natural language input (here, redundant cues) play an important role in learning, as indicated by performance in this lab-based learning task.

CONCLUSIONS AND FUTURE DIRECTIONS

As we hope has been reflected throughout this chapter, infants' accomplishments in the auditory domain are nothing short of extraordinary. In the absence of external guidance or reinforcement, our perceptual systems hone in on the dimensions of the auditory environment that are most relevant for the development of our communicative capacity, and we learn extremely complex and detailed information about how our auditory environment is structured, all during our 1st postnatal year. While much remains to be learned about how these

processes unfold, it is evident that they are heavily multidetermined, influenced by factors from the development of the peripheral auditory system to the nature of our learning mechanisms. We close by considering some limitations on these processes, which may be very important for future work aimed toward illuminating the nature of infants' accomplishments.

Relationship between Auditory Processing and Speech Perception

That even newborns can discriminate between speech sounds and recognize voices has led many to believe that hearing does not constrain speech perception or learning during infancy. It is clear, however, that several aspects of hearing remain immature early in infancy, and it is likely that these immaturities do constrain speech perception to some extent. There are suggestions in the literature that 2-month-olds, for example, represent speech with less detail than older infants do (Bertoncini et al., 1988; Bijeljac-Babic et al., 1993). It is likely that some aspects of speech perception and language learning are delayed until 6 months, when representations of the acoustic characteristics of sound are adultlike. In any case, it should be possible to make predictions about young infants' speech discrimination abilities based on what is known about their hearing, and to test specifically to determine whether hearing immaturity has any bearing on early speech perception.

A related question is whether infants use the same information in speech as adults do when they are discriminating between speech sounds. Because there are multiple cues to phonetic identity, it is possible that infants use cues that they hear better, or that they attend to more salient cues and ignore others, or that they weight all cues equally. That infants do not attend to the components of a complex sound as adults do in a simple psychophysical task (Bargones et al., 1995; Bargones & Werner, 1994; Leibold & Werner, 2003; Werner & Boike, 2001) suggests that their approach to speech may differ from that of adults. Nittrouer's studies of speech discrimination in children suggest that preschool children do not, in fact, weight cues to phonetic identity as adults do (e.g., Nittrouer, Crowther, & Miller, 1998; Nittrouer & Miller, 1997; Nittrouer & Studdert-Kennedy, 1987). It would be surprising to find, then, that infants weight cues in an adultlike way. There are now correlational techniques that can be used to assess the weights that listeners place on various components

of a complex sound in making discriminations and these techniques have been successfully applied to young children (e.g., Stellmack et al., 1997). An interesting problem in the future will be to apply these techniques to infants, particularly in the realm of speech perception.

Constraints on Learning

Much of the previous discussion has focused on infants' remarkable capacity to glean structure from complex input. However, it is important to note that demonstrations of powerful learning mechanisms alone do not represent a satisfying solution to the problems facing young language learners. How do learners hone in on the right patterns and structures given the massive amount of data in the input? The "richness of the stimulus problem" is that there are an infinite number of patterns that an unbiased learner might detect. Clearly, human infants are not such learners, and it is incumbent upon researchers to show not just all the things that infants can learn, but also what infants find more difficult to learn, to elucidate the limits on learning. It is also possible to ask how the structure of the task itself affects the types of learning that occur, as some types of input may elicit different learning mechanisms than others (e.g., Pena et al., 2003; Saffran, Reeck, Niehbur, & Wilson, 2005).

Thus far, this research strategy has primarily been carried out with adult learners, with implications to be drawn for infant learners. For example, Newport and Aslin (2004) demonstrated that while adults readily track the dependencies between adjacent syllables (e.g., the probability that *pa* is followed by *bu*), they do not do so when the relevant dependency skips an intervening syllable. Such nonadjacent dependencies are apparently not automatically tracked by learners. Interestingly, however, adults do detect nonadjacent dependencies when the intervening material is different in kind. For example, adults can detect dependencies between two consonants with intervening vowels, or two vowels with intervening consonants (Newport & Aslin, 2004). Because these latter types of structures recur in human languages (in Semitic languages, and in languages like Turkish that use vowel harmony), while the former do not, Newport and Aslin (2004) suggest that languages may be constrained by the limits on human learning. That is, only those structures that are learnable by humans persist in our languages. Saffran (2002) makes a similar argument based on adult grammatical studies.

The extent to which similar findings emerge with young learners is the object of active research

Domain Specificity and Species Specificity

Much of the foregoing discussion has focused on learning from the input, and the types of information captured by infant learning mechanisms. A critical open question is the degree to which this learning is subserved by mechanisms tailored for speech and language. One possibility is that, perhaps due to the adaptive significance of human communication systems, we have evolved sophisticated learning machinery specifically tailored for language. Alternatively, these early learning processes may tap mechanisms that are available for more general tasks.

A growing body of results suggests that at least one of the learning mechanisms we have discussed, sequential statistical learning, is quite general. For example, infants can track sequences of musical tones, discovering tone-word boundaries via statistical cues (e.g., Saffran, 2003a; Saffran & Griepentrog, 2001; Saffran, Johnson, Aslin, & Newport, 1999), and can learn statistically defined visual patterns (e.g., Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). These findings and others suggest that at least these basic learning processes are not tailored solely for language acquisition (e.g., Saffran, 2002, 2003b).

Another source of evidence bearing on this issue comes from studies of nonhuman primates. Hauser and his colleagues (Hauser, Newport, & Aslin, 2001; Hauser, Weiss, & Marcus, 2002) have tested cotton-top tamarins, a new world monkey species, on the linguistic tasks used by Saffran et al. (1996) and Marcus et al. (1999). Intriguingly, the monkeys showed the same pattern of performance as human infants, despite their presumed lack of evolved abilities to acquire human language (Hauser et al., 2001, 2002). Even rats detect some language-relevant patterns (Toro & Trobalan, 2004)! These findings reinforce the view that at least some of the learning mechanisms that subserve the beginnings of language learning are not evolutionary adaptations specialized for the linguistic domain.

Results like these lead immediately to the question of why, if monkeys share our learning machinery, language is uniquely human. That is, if monkeys learn like us, shouldn't they be as linguistically sophisticated as we are? Several avenues of explanation are currently being

94 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

explored. One, of course, is the traditional view that humans possess innate linguistic knowledge that other species lack (e.g., Pinker, 1984). Other investigators are focusing on the degree to which human learning mechanisms may in fact diverge from those possessed by other species. For example, tamarins and human adults do not show the same pattern of learning of nonadjacencies discussed in the previous section, suggesting that the constraints on human learning mechanisms may diverge from those seen in other species (Newport, Hauser, Spaepen, & Aslin, 2004). Similarly, Hauser, Chomsky, and Fitch (2002) have suggested that while humans and nonhumans may share much of their learning machinery, humans are differentiated by their ability to perform recursion operations—the capacity to generate an infinite range of expressions from a finite set of elements (see also Fitch & Hauser, 2004). On this view, humans and nonhumans should show similar performance when learning about such things as speech contrasts and word segmentation, and diverge as grammatical complexity increases (e.g., Saffran, Hauser, Seibel, Kapfhamer, Tsao, & Cushman, 2005). While these central questions remain to be resolved, their answers are likely to have broad impact on such issues as the modularity of mind and the ontogenesis of specific domains of knowledge.

There is another component to early language learning that may be relevant to species differences—social interaction between the speaker and the learner. While such issues as joint attention have played a prominent role in the literature on how young children map sound to meaning (e.g., Harris, Chapter 19; Tomasello, Chapter 6, this *Handbook*, this volume), the role of social interaction has not received significant attention in the literature on how infants acquire sound structure itself. Certainly, there is ample evidence that caregivers manipulate the input so it is well tailored to infants' perceptual predilections (e.g., Kuhl et al., 1997; see Trehub, 2003, for related evidence in the domain of music perception). The higher pitches and enhanced pitch contours of infant directed speech are well-established attention-getters and affect communicators (e.g., Cooper & Aslin, 1990; Fernald, 1992). Intriguing new results suggest, however, that the role of social interaction extends beyond the sound structure of the input. Kuhl, Tsao, and Liu (2003) manipulated infants' perception of nonnative speech contrasts such that English-learning infants maintained a Mandarin speech contrast well beyond the age at which their ability to discriminate the contrast would typically have declined. Critically, how-

ever, human social interaction was required in the presentation of the Mandarin input. When infants received the same input via high-quality DVD recordings, no impact on their speech perception was observed. These results suggest that, like some species of birds, the learning system requires a certain type of interactive input to affect perception. If this is the case, then differences in social interaction may also help to explain some cross-species differences in who learns what.

The Infant's Auditory World

In this review, we have considered recent developments in our understanding of how infants begin to make sense of their auditory environments. A great deal of progress has been made in elucidating the basic sensory and perceptual mechanisms that provide auditory input to infant learners, as well as the learning mechanisms that track this input and integrate it with infants' existing knowledge.

In future studies, we expect that the relationship between infants' auditory abilities and the rest of language acquisition (see Tomasello, Chapter 6; Waxman & Lidz, Chapter 7, this *Handbook*, this volume) will become clearer. Audition is the gateway to spoken language, and infants' early accomplishments in acquiring the sound structure of their native language(s) lay critical groundwork for subsequent learning. Recent studies linking the acquisition of sound structure to later accomplishments in word learning provide important suggestions about how infants' early abilities are likely to influence later language learning (e.g., Hollich, Jusczyk, & Luce, 2002; Saffran & Graf Estes, 2004; Swingley & Aslin, 2002; Thiessen, 2004; Werker et al., 2002). For example, early speech perception abilities may predict some aspects of word learning many months later (Tsao, Liu, & Kuhl, 2004). Similarly, researchers are beginning to investigate the effects of the amelioration of early sensory deprivation via cochlear implants on subsequent auditory perception and language learning abilities (Houston, Pisoni, Kirk, Ying, & Miyamoto, in press). Such integrative research enterprises will serve to illuminate the links between the talents of infant listeners in the auditory realm and the many linguistic (and nonlinguistic) tasks that lie ahead of them. Similarly, much remains to be learned about the neural underpinnings of the abilities described throughout our review, and knowledge about these neural substrates will help us to better understand the behaviors that they subservise. Many fasci-

nating open questions thus remain, and in the next edition of the Handbook, we hope to read the answers—including the answers to the many questions that we do not yet know to ask.

REFERENCES

- Abdala, C. (1998). A developmental study of distortion product otoacoustic emission (2f₁-f₂) suppression in humans. *Hearing Research, 121*, 125–138.
- 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., & 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., & Folsom, R. C. (1995a). 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*, 921–930.
- Abdala, C., & Folsom, R. C. (1995b). Frequency contribution to the click-evoked auditory brain stem response in human adults and infants. *Journal of the Acoustical Society of America, 97*, 2394–2404.
- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, Scotland: Edinburgh University Press.
- 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*, 2997–3005.
- Allen, P., & Wightman, F. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research, 35*, 222–233.
- 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.
- Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research, 32*, 317–322.
- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.
- Altmann, G. T. M., & Dienes, Z. (1999). Rule learning by 7-month-old infants and neural networks. *Science, 284*, 875.
- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*(2/3), 155–182.
- Ashmead, D. H., Clifton, R. K., & Perris, E. E. (1987). Precision of auditory localization in human infants. *Developmental Psychology, 23*, 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*, 1211–1226.
- Aslin, R. N. (1989). Discrimination of frequency transitions by human infants. *Journal of the Acoustical Society of America, 86*, 582–590.
- Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception* (pp. 67–96). New York: Academic Press.
- Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Percy, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development, 52*(4), 1135–1145.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321–324.
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134). Hillsdale, NJ: Erlbaum.
- 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*, 1809–1816.
- 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.
- Batchelder, E. O. (1997). *Computational evidence for the use of frequency information in discovery of the infant's first lexicon*. Unpublished doctoral dissertation, New York, City University.
- Beckman, M. E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development, 71*(1), 240–249.
- Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., et al. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language, 78*, 364–396.
- 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*, 183–190.
- Berg, K. M., & Boswell, A. E. (1998). Infants' detection of increments in low- and high-frequency noise. *Perception and Psychophysics, 60*, 1044–1051.
- Berg, K. M., & Boswell, A. E. (1999). Effect of masker level on infants' detection of tones in noise. *Perception and Psychophysics, 61*, 80–86.
- Berg, K. M., & Boswell, A. E. (2000). Noise increment detection in children 1 to 3 years of age. *Perception and Psychophysics, 62*, 868–873.
- Berg, K. M., & Smith, M. C. (1983). Behavioral thresholds for tones during infancy. *Journal of Experimental Child Psychology, 35*, 409–425.
- Bertoncini, J., Bijeljac-Babic, R., Blumstein, S., & Mehler, J. (1987). Discrimination of very short CV syllables by neonates. *Journal of the Acoustical Society of America, 82*, 31–37.
- 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., & Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behavior and Development, 4*(3), 247–260.
- Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., & Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain and Language, 37*(4), 591–605.

96 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.
- Best, C. T., Hoffman, H., & Glanville, B. B. (1982). Development of infant ear asymmetries for speech and music. *Perception and Psychophysics*, 31(1), 75–85.
- Best, C. T., & McRoberts, G. W. (2003). Infant perception of nonnative contrasts that adults assimilate in different ways. *Language and Speech*, 46(2/3), 183–216.
- Best, C. T., McRoberts, G. W., LaFleur, R., & Silver Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, 18(3), 339–350.
- 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(3), 345–360.
- Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, 29(4), 711–721.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10(5), 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17(1), 353–362.
- Binns, K. E., Withington, D. J., & Keating, M. J. (1995). The developmental emergence of the representation of auditory azimuth in the external nucleus of the inferior colliculus of the guinea-pig: The effects of visual and auditory deprivation. *Developmental Brain Research*, 85, 14–24.
- Bond, B., & Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception and Psychophysics*, 6, 337–339.
- Bonfils, P., Avan, P., Francois, M., Trotoux, J., & Narcy, P. (1992). Distortion-product otoacoustic emissions in neonates: Normative data. *Acta Otolaryngologica*, 112, 739–744.
- Bonfils, P., Francois, M., Avan, P., Londero, A., Trotoux, J., & Narcy, P. (1992). Spontaneous and evoked otoacoustic emissions in preterm neonates. *Laryngoscope*, 102, 182–186.
- Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). *Mommy and me: Familiar names help launch babies into speech stream segmentation*. Manuscript in preparation.
- Bosch, L., & Sebastián-Gallés, N. (2001). Early language differentiation in bilingual infants. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 71–93). Amsterdam: Benjamins.
- Bredberg, G. (1968). Cellular pattern and nerve supply of the human organ of Corti. *Acta Otolaryngologica* (Suppl.), 236.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61(1/2), 93–125.
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44.
- Brown, A. M., Sheppard, S. L., & 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, C., & Matthews, J. (1997). The role of feature geometry in the development of phonemic contrasts. In S. J. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition: Vol. 16. Language acquisition and language disorders* (pp. 67–112). Amsterdam: Benjamins.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Bull, D., Eilers, R. J., & Oller, D. K. (1984). Infants' discrimination of intensity variation in multisyllabic stimuli. *Journal of the Acoustical Society of America*, 76(1), 13–17.
- Burnham, D., & Dodd, B. (2004). Audiovisual speech perception by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, 45(4), 202–220.
- Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science*, 296(5572), 1435.
- 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, 384–394.
- 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, 844–849.
- Cairns, P., Shillcock, R., Chater, N., & Levy, J. (1997). Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology*, 33(2), 111–153.
- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experiences. *Cognition*, 87(2), B69–B77.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., et al. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1(5), 351–353.
- Cheour-Luhtanen, M., Alho, K. I., Kujala, T., Sainio, K., Reinikainen, K., et al. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research*, 82(1), 53–58.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 2–3.
- Christiansen, M. H., & Curtin, S. L. (1999). Transfer of learning: Rule acquisition or statistical learning? *Trends in Cognitive Sciences*, 3, 289–290.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, 95(3), 1570–1580.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). *Discovering words in the continuous speech stream: The role of prosody*. Manuscript in preparation.
- Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, 2, 358–394.
- Church, K. W. (1987). Phonological parsing and lexical retrieval. *Cognition*, 25, 53–69.
- Church, B. A., & Fisher, C. (1998). Long-term auditory word priming in preschoolers: Implicit memory support for language acquisition. *Journal of Memory and Language*, 39, 523–542.

- 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., & Clifton, R. K. (1995). Infants' pitch perception: Inharmonic tonal complexes. *Journal of the Acoustical Society of America*, *98*(3), 1372–1379.
- Clarkson, M. G., Clifton, R. K., & Perris, E. E. (1988). Infant timbre perception: Discrimination of spectral envelopes. *Perception and Psychophysics*, *43*, 15–20.
- 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.
- 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.
- Clifton, R. K., Morrongiello, B. A., & Dowd, J. M. (1984). A developmental look at an auditory illusion: The precedence effect. *Developmental Psychobiology*, *17*, 519–536.
- Clifton, R. K., Morrongiello, B. A., Kulig, J. W., & Dowd, J. M. (1981). Developmental changes in auditory localization in infancy. In R. Aslin, J. Alberts, & M. R. Petersen (Eds.), *Development of perception* (Vol. 2, pp. 141–160). New York: Academic Press.
- Clifton, R. K., Perris, E. E., & Bullinger, A. (1991). Infants' perception of auditory space. *Developmental Psychology*, *27*, 187–197.
- Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, *30*, 441–469.
- Coady, J. A., & Aslin, R. N. (in press). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology*.
- Cole, R., & Jakimik, J. (1980). *A model of speech perception*. Hillsdale, NJ: Erlbaum.
- 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*, 2012–2021.
- Colombo, J. A., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, *4*(2), 219–223.
- Colombo, J., & Horowitz, F. D. (1986). Infants' attentional responses to frequency modulated sweeps. *Child Development*, *57*(2), 287–291.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595.
- Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, *65*(6), 1663–1677.
- Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*, 233–262.
- Curtin, S., Werker, J. F., & Ladhar, N. (2002). Accessing coarticulatory information. *Journal of the Acoustical Society of America*, *112*, 2359.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 3–4.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature*, *304*(5922), 159–160.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*(4), 385–400.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113–121.
- Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*, *128*(2), 165–185.
- 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.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176.
- DeCasper, A. J., & Prescott, P. (1984). Human newborns' perception of male voices: Preference, discrimination and reinforcing value. *Developmental Psychobiology*, *17*, 481–491.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*, 133–150.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, D., Perani, D., et al. (1997). Anatomical variability in the cortical representation of first and second languages. *Neuroreport: For Rapid Communication of Neuroscience Research*, *8*, 3809–3815.
- Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport*, *9*(8), 1885–1888.
- Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, *370*(6487), 292–295.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600), 2013–2015.
- Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *NeuroReport*, *12*(14), 3155–3158.
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1355–1367.
- Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development*, *5*, 261–276.
- Demany, L. (1985). Perceptual learning in frequency discrimination. *Journal of the Acoustical Society of America*, *78*, 1118–1120.
- Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Science*, *266*, 718–719.
- Desjardins, R. N., Rogers, J., & Werker, J. F. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology*, *66*(1), 85–110.
- Desjardins, R., & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology*, *45*(4), 187–203.
- Dewson, J. H. (1964). Speech sound discrimination by cats. *Science*, *141*, 555–556.
- Dooling, R. J., Best, C. T., & Brown, S. D. (1995). Discrimination of synthetic full-formant and sinewave/rala/continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). *Journal of the Acoustical Society of America*, *97*(3), 1839–1846.
- Durieux-Smith, A., Edwards, C. G., Picton, T. W., & McMurray, B. (1985). Auditory brainstem responses to clicks in neonates. *Journal of Otolaryngology*, *14*, 12–18.

98 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, 36, 202–225.
- Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first words. *Language Acquisition*, 2(3), 189–220.
- Eddins, D. A., & Green, D. M. (1995). Temporal integration and temporal resolution. In B. C. J. Moore (Ed.), *Hearing* (pp. 207–242). San Diego, CA: Academic Press.
- Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perceptual Psychophysiology*, 16, 513–521.
- Eimas, P. D. (1975a). Auditory and phonetic coding of the cues for speech: Discrimination of the {r-l} distinction by young infants. *Perception and Psychophysics*, 18(5), 341–347.
- Eimas, P. D. (1975b). Speech perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (pp. 193–231). New York: Academic Press.
- Eimas, P. D., & Miller, J. L. (1981). Organization in the perception of segmental and suprasegmental information by infants. *Infant Behavior and Development*, 4, 395–399.
- Eimas, P. D., & Miller, J. L. (1992). Organization in the perception of speech by young infants. *Psychological Science*, 3, 340–345.
- Eimas, P. D., Miller, J. L., & Jusczyk, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 161–195). New York: Cambridge University Press.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(968), 303–306.
- Elfenbein, J. L., Small, A. M., & Davis, M. (1993). Developmental patterns of duration discrimination. *Journal of Speech and Hearing Research*, 36, 842–849.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Fassbender, C. (1993). *Auditory grouping and segregation processes in infancy*. Norderstedt, Germany: Kaste Verlag.
- Fay, R. R., & Coombs, S. (1983). Neural mechanisms in sound detection and temporal summation. *Hearing Research*, 10, 69–92.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46(2/3), 245–264.
- Fernald, A. (Ed.). (1984). *The perceptual and affective salience of mothers' speech to infants*. Norwood, NJ: Ablex.
- Fernald, A. (1992). Prosody in speech to children: Prelinguistic and linguistic functions. *Annals of Child Development*, 8.
- 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(3), 228–231.
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, 72(4), 1003–1015.
- Fiez, J. A., Tallal, P. A., Raichle, M. E., Miezin, F. M., Katz, W., Dohmeyer, S., et al. (1995). PET studies of auditory and phonological processing: Effects of stimulus type and task condition. *Journal of Cognitive Neuroscience*, 7, 357–375.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, 99, 15822–15826.
- Fisher, C. H., Hunt, C. M., Chambers, C. K., & Church, B. (2001). Abstraction and specificity in preschoolers' representations of novel spoken words. *Journal of Memory and Language*, 45(4), 665–687.
- Fisher, C., Church, B., & Chambers, K. E. (2004). Learning to identify spoken words. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 3–40). Cambridge, MA: MIT Press.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67(6), 3192–3218.
- Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303, 377–380.
- Floccia, C., Nazzi, T., & Bertoncini, J. (2000). Unfamiliar voice discrimination for short stimuli in newborns. *Developmental Science*, 3(3), 333–343.
- Folsom, R. C., & Wynne, M. K. (1987). Auditory brain stem responses from human adults and infants: Wave 5—Tuning curves. *Journal of the Acoustical Society of America*, 81, 412–417.
- Fowler, C. A., Smith, M. R., & Tassinari, L. G. (1986). Perception of syllable timing by prebabbling infants. *Journal of the Acoustical Society of America*, 79, 814–825.
- Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics*, 54(3), 287–295.
- Fujikawa, S. M., & Weber, B. A. (1977). Effects of increased stimulus rate on brainstem electric response (BER) audiometry as a function of age. *Journal of the American Audiology Society*, 3, 147–150.
- Fujimoto, S., Yamamoto, K., Hayabuchi, I., & Yoshizuka, M. (1981). Scanning and transmission electron microscope studies on the organ of corti and stria vascularis in human fetal cochlear ducts. *Archives of Histology Japan*, 44, 223–235.
- Fujita, A., Hyde, M. L., & Alberti, P. W. (1991). ABR latency in infants: Properties and applications of various measures. *Acta Otolaryngologica*, 111, 53–60.
- Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language*, 16, 161–179.
- Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition*, 93, B89–B96.
- Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, 51(3), 237–265.
- Gerken, L. A., & McIntosh, B. J. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29, 448–457.
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). 17-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268.
- Ghazanfar, A. A., & Logothetis, N. K. (2003). Facial expressions linked to monkey calls. *Nature*, 423, 937–938.
- Glanville, B. B., Levenson, R., & Best, C. T. (1977). A cardiac measure of cerebral asymmetries in infant auditory perception. *Developmental Psychology*, 13(1), 54–49.
- Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. R. Wanner (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge, England: Cambridge University Press.
- Glenn, S. M., Cunningham, C. C., & Joyce, P. F. (1981). A study of auditory preferences on nonhandicapped infants with Down's syndrome. *Child Development*, 52, 1303–1307.
- Goldinger, S. D. (1992). Words and voices: Implicit and explicit memory for spoken words. *Dissertation Abstracts International*, 53(6), 3189.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431–436.

- Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
- Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4, 178–186.
- Gómez, R. L., & Maye, J. (2005). The developmental trajectory of non-adjacent dependency learning. *Infancy*, 7, 183–206.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, 20, 229–252.
- Gorga, M. P., Kaminski, J. R., Beauchaine, K. L., Jesteadt, W., & Neely, S. T. (1989). Auditory brainstem responses from children 3 months to 3 years of age: Vol. 2. Normal patterns of response. *Journal of Speech and Hearing Research*, 32, 281–288.
- Gorga, M. P., Reiland, J. K., Beauchaine, K. A., Worthington, D. W., & Jesteadt, W. (1987). Auditory brainstem responses from graduates of an intensive care nursery: Normal patterns of response. *Journal of Speech and Hearing Research*, 30, 311–318.
- Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Neural and behavioral specificity* (pp. 25–53). New York: Academic Press.
- Gray, L. (1992). Interactions between sensory and nonsensory factors in the responses of newborn birds to sound. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 89–112). Washington, DC: American Psychological Association.
- Green, K. P. (1998). The use of auditory and visual information during phonetic processing: Implications for theories of speech perception. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye: Vol. 2. Advances in the psychology of speechreading and auditory-visual speech* (pp. 3–25). Hove, England: Psychology Press.
- Greenberg, G. Z., Bray, N. W., & Beasley, D. S. (1970). Children's frequency-selective detection of signals in noise. *Perception and Psychophysics*, 8, 173–175.
- Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, 25(4), 577–588.
- Grose, J. H., Hall, J. W. I., & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech and Hearing Research*, 36, 351–356.
- Hall, J. W., & Grose, J. H. (1990). The masking level difference in children. *Journal of the American Academy of Audiology*, 1, 81–88.
- 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, 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, 150–154.
- 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.
- Hallé, P. A., & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development*, 17, 119–129.
- Harris, Z. (1955). From phoneme to morpheme. *Language*, 31, 190–222.
- Hartley, D. E. H., & Moore, D. R. (2002). Auditory processing efficiency deficits in children with developmental language impairments. *Journal of the Acoustical Society of America*, 112, 2962–2966.
- Hartley, D. E. H., 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, 1402–1415.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B53–B64.
- Hauser, M. D., Weiss, D., & Marcus, G. (2002). Rule learning by cotton-top tamarins. *Cognition*, 86, B15–B22.
- Hayes, R. A., Slater, A., & Brown, E. (2000). Infants' ability to categorize on the basis of rhyme. *Cognitive Development*, 15(4), 405–419.
- Hillenbrand, J. A. (1984). Speech perception by infants: Categorization based on nasal consonant place of articulation. *Journal of the Acoustical Society of America*, 75(950), 1613–1622.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, 26, 269–286.
- Hockley, N. S., & Polka, L. (1994). *A developmental study of Audiovisual Speech Perception Using the McGurk Paradigm*. Unpublished masters thesis, McGill University, Montreal, Canada.
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, 6(2), 122–127.
- Hohne, E. A., & Jusczyk, P. W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics*, 56, 613–623.
- Hollich, G., Jusczyk, P., & Luce, P. (2002). Lexical neighborhood effects in 17-month-old word learning. *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 314–323). Boston: Cascadilla Press.
- Hoshino, T. (1990). Scanning electron microscopy of nerve fibers in human fetal cochlea. *Journal of Electron Microscopy Technique*, 15, 104–114.
- 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(5), 1570–1582.
- Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, 7(3), 504–509.
- Houston, D. M., Pisoni, D. B., Kirk, K. I., Ying, E. A., & Miyamoto, R. T. (in press). Speech perception skills of deaf infants following cochlear implantation: A first report. *International Journal of Pediatric Otorhinolaryngology*.
- Houston, D. M., Santelmann, L., & Jusczyk, P. W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, 19, 97–136.
- Igarashi, Y., & Ishii, T. (1979). Development of the cochlear and the blood-vessel-network in the fetus: A transmission electrographic observation. *Audiology, Japan*, 22, 459–460.
- Igarashi, Y., & Ishii, T. (1980). Embryonic development of the human organ of Corti: Electron microscopic study. *International Journal of Pediatric Otorhinolaryngology*, 2, 51–62.
- Igarashi, Y., Yamazaki, H., & Mitsui, T. (1978). An electronographic study of inner/outer haircells of human fetuses. *Audiology, Japan*, 21, 375–377.
- Ilari, B., & Polka, L. (2002). *Memory for music in infancy: The role of style and complexity*. Paper presented at the International Conference on Infant Studies, Toronto, Ontario, Canada.

100 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Irwin, R. J., Stillman, J. A., & Schade, A. (1986). The width of the auditory filter in children. *Journal of Experimental Child Psychology*, *41*, 429–442.
- Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, *4*, 104–107.
- Jiang, Z. D., Wu, Y. Y., & Zhang, L. (1990). Amplitude change with click rate in human brainstem auditory-evoked responses. *Audiology*, *30*, 173–182.
- Jiang, Z. D., Wu, Y. Y., Zheng, W. S., Sun, D. K., Feng, L. Y., & Liu, X. Y. (1991). The effect of click rate on latency and interpeak interval of the brain-stem auditory evoked potentials in children from birth to 6 years. *Electroencephalography and Clinical Neurophysiology*, *80*, 60–64.
- 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*(4), 548–567.
- Johnson, E. K., Jusczyk, P. W., Cutler, A., & Norris, D. (2003). Lexical viability constraints on speech segmentation by infants. *Cognitive Psychology*, *46*(1), 65–97.
- Jusczyk, P. W. (1997). Finding and remembering words: Some beginnings by English-learning infants. *Current Directions in Psychological Science*, *6*(6), 170–174.
- 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., Copan, H., & Thompson, E. (1978). Perception by 2-month-old infants of glide contrasts in multisyllabic utterances. *Perception and Psychophysics*, *24*(6), 515–520.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675–687.
- Jusczyk, P. W., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, *23*, 648–654.
- Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science*, *277*, 1984–1986.
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*, 402–420.
- Jusczyk, P. W., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*, *40*(1), 62–82.
- Jusczyk, P. W., Hirsh-Pasek, K., Nelson, D. G. K., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.
- Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues to word segmentation. *Perception and Psychophysics*, *61*, 1465–1476.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3/4), 159–207.
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(3), 627–640.
- 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*(5), 630–645.
- Jusczyk, P. W., Pisoni, D. B., Reed, M. A., Fernald, A., & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. *Science*, *222*, 175–177.
- Jusczyk, P. W., Pisoni, D. B., Walley, A., & Murray, J. (1980). Discrimination of relative onset time of two-component tones by infants. *Journal of the Acoustical Society of America*, *67*(1), 262–270.
- Kaga, K., & Tanaka, Y. (1980). Auditory brainstem response and behavioral audiometry: Developmental correlates. *Archives of Otolaryngology*, *106*, 564–566.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Karzon, R. G. (1985). Discrimination of polysyllabic sequences by 1- to 4-month-old infants. *Journal of Experimental Child Psychology*, *39*, 326–342.
- Karzon, R. G., & Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics*, *45*(1), 10–14.
- 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.
- Keefe, D. H., Folsom, R. C., Gorga, M. P., Vohr, B. R., Bulen, J. C., & Norton, S. J. (2000). Identification of neonatal hearing impairment: Ear-canal measurements of acoustic admittance and reflectance in neonates. *Ear and Hearing*, *21*, 443–461.
- Keefe, D. H., & Levi, E. C. (1996). Maturation of the middle and external ears: Acoustic power-based responses and reflectance tympanometry. *Ear and Hearing*, *17*, 1–13.
- Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, *99*(2), 349–364.
- Kidd, G., Jr., 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.
- King, A. J., Hutchings, M. E., Moore, D. R., & Blakemore, C. (1988). Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature*, *332*, 73–76.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42.
- Klein, A. J., Alvarez, E. D., & Cowburn, C. A. (1992). The effects of stimulus rate on detectability of the auditory brain stem response in infants. *Ear and Hearing*, *13*, 401–405.
- Kluender, K. R., Diehl, R. L., & Killeen, P. R. (1987, September 4). Japanese quail can learn phonetic categories. *Science*, *237*, 1195–1197.
- Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America*, *104*, 3568–3582.
- Kopyar, B. A. (1997). *Intensity discrimination abilities of infants and adults: Implications for underlying processes*. Unpublished doctoral dissertation, University of Washington, Seattle.
- Krentz, U. C., & Corina, D. C. (2005). *Preference for language in early infancy: The human language bias is not speech specific*. Manuscript submitted for publication.
- Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, *1*(1), 70–73.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66*(6), 1668–1679.

- Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*(2), 93–107.
- Kuhl, P. K. (2000). Language, mind, and brain: Experience alters perception. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 99–115). Cambridge, MA: MIT Press.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 684–686.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141.
- Kuhl, P. K., & Meltzoff, A. N. (1988). Speech as an intermodal object of perception. In A. Yonas (Ed.), *Minnesota Symposia on Child Psychology: Perceptual development in infancy* (pp. 235–266). Hillsdale, NJ: Erlbaum.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chin-chilla: Identification function for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, *63*, 905–917.
- Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, *73*, 1003–1010.
- 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, USA*, *100*, 9096–9101.
- Kuhl, P. K., Tsao, F.-M., Liu, H.-M., Zhang, Y., & de Boer, B. (2001). Language/Culture/Mind/Brain: Progress at the margins between disciplines. In A. Domasio et al. (Eds.), *Unity of knowledge: The convergence of natural and human science* (pp. 136–174). New York: New York Academy of Sciences.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.
- Lary, S., Briassoulis, G., de Vries, L., Dubowitz, L. M. S., & Dubowitz, V. (1985). Hearing threshold in preterm and term infants by auditory brainstem response. *Journal of Pediatrics*, *107*, 593–599.
- Lasky, R. E. (1984). A developmental study on the effect of stimulus rate on the auditory evoked brain-stem response. *Electroencephalography and Clinical Neurophysiology*, *59*, 411–419.
- Lasky, R. E. (1991). The effects of rate and forward masking on human adult and newborn auditory evoked response thresholds. *Developmental Psychobiology*, *24*, 21–64.
- Lasky, R. E. (1993). The effect of forward masker duration, rise/fall time, and integrated pressure on auditory brain stem evoked responses in human newborns and adults. *Ear and Hearing*, *14*, 95–103.
- Lasky, R. E. (1997). Rate and adaptation effects on the auditory evoked brainstem response in human newborns and adults. *Hearing Research*, *111*, 165–176.
- Lasky, R. E., & Rupert, A. (1982). Temporal masking of auditory evoked brainstem responses in human newborns and adults. *Hearing Research*, *6*, 315–334.
- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). Vot discrimination by 4- to 6-month-old infants from Spanish environments. *Journal of Experimental Child Psychology*, *20*, 215–225.
- Lavigne-Rebillard, M., & Bagger-Sjoberg, D. (1992). Development of the human stria vascularis. *Hearing Research*, *64*, 39–51.
- Lavigne-Rebillard, M., & Pujol, R. (1987). Surface aspects of the developing human organ of corti. *Acta Otolaryngologica*, *436*, 43–50.
- Lavigne-Rebillard, M., & Pujol, R. (1988). Hair cell innervation in the fetal human cochlea. *Acta Otolaryngologica, Stockholm*, *105*, 398–402.
- Lavigne-Rebillard, M., & Pujol, R. (1990). Auditory hair cells in human fetuses: Synaptogenesis and ciliogenesis. *Journal of Electron Microscopy Technique*, *15*, 115–122.
- Lecanuet, J.-P., Granier-Deferre, C., & Busnel, M.-C. (1988). Fetal cardiac and motor responses to octave-band noises as a function of central frequency, intensity and heart rate variability. *Early Human Development*, *18*, 81–93.
- Leibold, L., & Werner, L. A. (2002). Relationship between intensity and reaction time in normal hearing infants and adults. *Ear and Hearing*, *23*, 92–97.
- Leibold, L., & Werner, L. A. (2003). Infants' detection in the presence of masker uncertainty. *Journal of the Acoustical Society of America*, *113*, 2208.
- Levi, E. C., & Werner, L. A. (1996). Amplitude modulation detection in infancy: Update on 3-month-olds. *Abstracts of the Association for Research in Otolaryngology*, *19*, 142.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.
- Lisker, L., & Abramson, A. S. (1971). Distinctive features and laryngeal control. *Language*, *47*, 767–785.
- Litovsky, R. Y. (1997). Developmental changes in the precedence effect: Estimates of minimum audible angle. *Journal of the Acoustical Society of America*, *102*(3), 1739–1745.
- Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1–F10.
- Lotto, A. J., Kluender, K. R., & Holt, L. L. (1998). Depolarizing the perceptual magnet effect. *Journal of the Acoustical Society of America*, *103*, 3648–3655.
- MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception process. *Perception and Psychophysics*, *24*, 253–257.
- MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983, March 18). Infant intermodal speech perception is a left-hemisphere function. *Science*, *219*, 1347–1349.
- Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53*, 155–180.
- Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, *6*(5), 315–318.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by 7-month-old infants. *Science*, *283*(5398), 77–80.
- Massaro, D. W., Cohen, M. M., & Smeele, P. M. T. (1995). Cross-linguistic comparisons in the integration of visual and auditory speech. *Memory and Cognition*, *23*(1), 113–131.
- Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology*, *41*, 93–113.
- Mattingsly, I. G., Liberman, A. M., Syrdal, A. K., & Halwes, T. (1971). Discrimination in speech and non-speech modes. *Cognitive Psychology*, *2*, 131–157.
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *78*(2), 91–121.
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–494.

102 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing, 3*, 301–308.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101–B111.
- McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal of the Acoustical Society of America, 102*(5, Pt. 1), 2945–2953.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.
- McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition, 95*(2), B15–B26.
- Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 943–954). Cambridge, MA: MIT Press.
- Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Cognitive Psychology, 18*, 1–86.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143–178.
- Miller, J. L. (1987). Rate-dependent processing in speech perception. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 119–157). Hillsdale, NJ: Erlbaum.
- Miller, J. L., & Eimas, P. D. (1979). Organization in infant speech perception. *Canadian Journal of Psychology, 33*, 353–367.
- Miller, J. L., & Eimas, P. D. (1996). Internal structure of voicing categories in early infancy. *Perception and Psychophysics, 58*(8), 1157–1167.
- Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience, 16*(8), 1–13.
- Moffitt, A. R. (1971). Consonant cue perception by 20- to 24-week-old infants. *Child Development, 42*, 505–511.
- Molfese, D. L., Burger-Judisch, L. M., & Hans, L. L. (1991). Consonant discrimination by newborn infants: Electrophysiological differences. *Developmental Neuropsychology, 7*(2), 177–195.
- Molfese, D. L., & Molfese, V. (1979). Hemisphere and stimulus differences as reflected in the cortical responses of newborn infants. *Developmental Psychology, 15*, 505–511.
- Molfese, D. L., & Molfese, V. (1980). Cortical responses of preterm infants to phonetic and nonphonetic speech stimuli. *Developmental Psychology, 16*, 574–581.
- Molfese, D. L., & Molfese, V. (1985). Electrophysiological indices of auditory discrimination in newborn infants: The bases for predicting later language development? *Infant Behavior and Development, 8*, 197–211.
- Montgomery, C. R., & Clarkson, M. G. (1997). Infants' pitch perception: Masking by low- and high-frequency noises. *Journal of the Acoustical Society of America, 102*, 3665–3672.
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development, 16*(4), 495–500.
- Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America, 54*, 610–619.
- Moore, B. C. J. (1996). *Introduction to the psychology of hearing* (5th ed.). New York: Academic Press.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., & Biswas, A. K. (1988). The shape of the ear's temporal window. *Journal of the Acoustical Society of America, 83*, 1102–1116.
- Moore, J. K. (2002). Maturation of human auditory cortex: Implications for speech perception. *Annals of Otology Rhinology and Laryngology, 111*, 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., 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., Perazzo, L. M., & Braun, A. (1995). Time course of axonal myelination in 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 and Hearing, 17*, 411–418.
- Mora, J. A., Exposito, M., Solis, C., & Barajas, J. J. (1990). Filter effects and low stimulation rate on the middle-latency response in newborns. *Audiology, 29*, 329–335.
- Morais, J., Content, A., Cary, L., Mehler, J., & Segui, J. (1989). Syllabic segmentation and literacy. *Language and Cognitive Processes, 4*(1), 57–67.
- Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Hillsdale, NJ: Erlbaum.
- Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior, 20*, 67–85.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of intonational and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498–550.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development, 66*, 911–936.
- Morgan, J. L., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In J. L. M. K. Demuth (Ed.), *Signal to syntax* (pp. 263–283). Hillsdale, NJ: Erlbaum.
- Morronegiello, B. A. (1988). Infants' localization of sounds in the horizontal plane: Estimates of minimum audible angle. *Developmental Psychology, 24*, 8–13.
- Morronegiello, B. A., Fenwick, K., & Chance, G. (1990). Sound localization acuity in very young infants: An observer-based testing procedure. *Developmental Psychology, 26*, 75–84.
- Morronegiello, B. A., Fenwick, K. D., Hillier, L., & Chance, G. (1994). Sound localization in newborn human infants. *Developmental Psychobiology, 27*, 519–538.
- Morronegiello, B. A., Hewitt, K. L., & Gotowiec, A. (1991). Infants' discrimination of relative distance in the auditory modality: Approaching versus receding sound sources. *Infant Behavior and Development, 14*, 187–208.
- Morronegiello, B. A., & Rocca, P. T. (1987a). Infants' localization of sounds in the horizontal plane: Effects of auditory and visual cues. *Child Development, 58*, 918–927.

- Morrongiello, B. A., & Rocca, P. T. (1987b). Infants' localization of sounds in the median sagittal plane: Effects of signal frequency. *Journal of the Acoustical Society of America*, *82*, 900–905.
- Morrongiello, B. A., & Rocca, P. T. (1987c). Infants' localization of sounds in the median vertical plane: Estimates of minimal audible angle. *Journal of Experimental Child Psychology*, *43*, 181–193.
- Morrongiello, B. A., & Rocca, P. T. (1990). Infants' localization of sounds within hemifields: Estimates of minimum audible angle. *Child Development*, *61*, 1258–1270.
- Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology*, *44*, 413–426.
- Morse, P. A. (1972). The discrimination of speech and non-speech stimuli in early infancy. *Journal of Experimental Child Psychology*, *14*(3), 477–492.
- Morse, P. A., & Snowden, C. T. (1975). An investigation of categorical speech discrimination by rhesus monkeys. *Perception and Psychophysics*, *17*(1), 9–16.
- Myers, J., Jusczyk, P. W., Kemler Nelson, D. G., Charles-Luce, J., Woodward, A. L., & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language*, *23*(1), 1–30.
- Naigles, L. (1990). Children use syntax to learn verb meaning. *Journal of Child Language*, *17*, 357–374.
- Nakai, Y. (1970). An electron microscopic study of the human fetus cochlea. *Practica of Otolaryngology, Rhinology and Laryngology*, *32*, 257–267.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756–766.
- Nazzi, T., Dille, L., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. (in press). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*.
- Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, *21*, 779–784.
- 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), 1–19.
- Neff, D. L., & Callaghan, B. P. (1988). Effective properties of multi-component simultaneous maskers under conditions of uncertainty. *Journal of the Acoustical Society of America*, *83*, 1833–1838.
- Neff, D. L., & Green, D. M. (1987). Masking produced by spectral uncertainty with multicomponent maskers. *Perception and Psychophysics*, *41*, 409–415.
- Nespor, M., Guasti, M. T., & Christophe, A. (1996). Selecting word order: The rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin, Germany: Akademie Verlag.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance: II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*, 85–117.
- Nittrouer, S., Crowther, C. S., & Miller, M. E. (1998). The relative weighting of acoustic properties in the perception of s+stop clusters by children and adults. *Perception and Psychophysics*, *60*, 51–64.
- Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, *101*, 3353–3366.
- Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, *30*, 319–329.
- Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., Cone-Wesson, B., et al. (2000). Identification of neonatal hearing impairment: Summary and recommendations. *Ear and Hearing*, *21*, 529–535.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: Maturational sequence and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, *73*(4), 295–305.
- Nozza, R. J. (1987). The binaural masking level difference in infants and adults: Developmental change in binaural hearing. *Infant Behavior and Development*, *10*, 105–110.
- Nozza, R. J. (1995). Estimating the contribution of non-sensory factors to infant-adult differences in behavioral thresholds. *Hearing Research*, *91*, 72–78.
- Nozza, R. J., & Henson, A. M. (1999). Unmasked thresholds and minimum masking in infants and adults: Separating sensory from nonsensory contributions to infant-adult differences in behavioral thresholds. *Ear and Hearing*, *20*, 483–496.
- 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.
- Nygaard, L. C., & Pisoni, D. B. (1995). Speech perception: New directions in research and theory. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition: Vol. 11. Speech, language, and communication* (2nd ed., pp. 63–96). San Diego, CA: Academic Press.
- Oh, E. L., & Lutfi, R. A. (1999). Informational masking by everyday sounds. *Journal of the Acoustical Society of America*, *106*, 3521–3528.
- 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.
- 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.
- Olivier, D. C. (1968). *Stochastic grammars and language acquisition mechanisms*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- Olsho, L. W. (1984). Infant frequency discrimination. *Infant Behavior and Development*, *7*, 27–35.
- Olsho, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior and Development*, *7*, 27–35.
- Olsho, L. W., Koch, E. G., & Carter, E. A. (1988). Nonsensory factors in infant frequency discrimination. *Infant Behavior and Development*, *11*, 205–222.
- 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*, 1316–1324.
- 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., Schoon, C., Sakai, R., Turpin, R., & Sperduto, V. (1982). Auditory frequency discrimination in infancy. *Developmental Psychology*, *18*, 721–726.

104 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, 83(1), B13–B23.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32(2), 258–278.
- Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language*, 45(4), 526–545.
- Pater, J., Stager, C. L., & Werker, J. F. (2004). The lexical acquisition of phonological contrasts. *Language*, 80(3), 361–379.
- Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4½-month-old infants. *Infant Behavior and Development*, 22(2), 237–247.
- Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*, 81(1), 93–115.
- Patterson, M. L., & Werker, J. F. (2003). Two-month-olds match vowel information in the face and voice. *Developmental Science*, 6(2), 191–196.
- Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, 102(6), 3742–3753.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., et al. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100(20), 11702–11705.
- Phillips, C., Pellathy, T., & Marantz, A. (1999, October). *Magnetic mismatch field elicited by phonological feature contrast*. Paper presented at the Cognitive Neuroscience Society Meeting, Washington, DC.
- Pike, K. L. (1945). Step-by-step procedure for marking limited intonation with its related features of pause, stress and rhythm. In C. C. Fries (Ed.), *Teaching and learning English as a foreign language* (pp. 62–74). Ann Arbor, MI: Publication of the English Language Institute.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1989) *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Pinker, S. (1994). *The language instinct*. New York: HarperCollins.
- Pisoni, D. B. (1977). Identification and discrimination of the relative onset of two component tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, 61, 1352–1361.
- Plessinger, M. A., & Woods, J. R. (1987). Fetal auditory brain stem response: Effect of increasing stimulus rate during functional auditory development. *American Journal of Obstetrics and Gynecology*, 157, 1382.
- Polka, L., & Bohn, O. S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, 100(1), 577–592.
- Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, 109(Suppl. 5, Pt. 1), 2190–2201.
- Polka, L., Sundara, M., & Blue, S. (2002, June 3–7). *The role of language experience in word segmentation: A comparison of English, French, and bilingual infants*. Paper presented at the the 143rd Meeting of the Acoustical Society of America: Special Session in Memory of Peter Jusczyk, Pittsburgh, PA.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421–435.
- Pollack, I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America*, 57(Suppl. 1), 5.
- Pons, F. (in press). The effects of distributional learning on rats' sensitivity to phonetic information. *Journal of Experimental Psychology: Animal Behavior Processes*.
- 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, 1576–1587.
- Ponton, C. W., Eggermont, J. J., Don, M., Waring, M. D., Kwong, B., Cunningham, J., et al. (2000). Maturation of the mismatch negativity: Effects of profound deafness and cochlear implant use. *Audiology and Neuro Otology*, 5, 167–185.
- Ponton, C. W., Moore, J. K., & Eggermont, J. J. (1996). Auditory brain stem response generation by parallel pathways: Differential maturation of axonal conduction time and synaptic transmission. *Ear and Hearing*, 17, 402–410.
- Price, C., Wise, R., Warburton, E., Moore, C., Howard, D., Patterson, K., et al. (1996). Hearing and saying: The functional neuroanatomy of auditory word processing. *Brain*, 119(3), 919–931.
- Pujol, R., & Lavigne-Rebillard, M. (1992). Development of neurosensory structures in the human cochlea. *Acta Otolaryngologica*, 112, 259–264.
- Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2, 85–115.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288(5464), 349–351.
- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America*, 105(1), 512–521.
- Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292.
- Ratner, N. B. (1984). Patterns of vowel modification in mother-child speech. *Journal of Child Language*, 11, 557–578.
- Ratner, N. B., & Luberoff, A. (1984). Cues to post-vocalic voicing in mother-child speech. *Journal of Phonetics*, 12, 285–289.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, 212(4497), 947–949.
- Repp, B. H. (1984). Against a role of "chirp" identification in duplex perception. *Perception and Psychophysics*, 35(1), 89–93.
- Rice, K., & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental acquisition. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 23–42). Hillsdale, NJ: Erlbaum.
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8, 162–172.
- Rosen, S., van der Lely, H., Adlard, A., & Manganari, E. (2000). Backward masking in children with and without language disorders [Abstract]. *British Journal of Audiology*, 34, 124.

- Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception and Psychophysics*, 59(3), 347–357.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Ruth, R. A., Horner, J. S., McCoy, G. S., & Chandler, C. R. (1983). Comparison of auditory brainstem response and behavioral audiometry in infants. *Audiology, Scandinavian*, 17, 94–98.
- Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, 44(4), 493–515.
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, 47(1), 172–196.
- Saffran, J. R. (2003a). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, 6(1), 35–43.
- Saffran, J. R. (2003b). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110–114.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- Saffran, J. R., & Graf Estes, K. M. (2004, May). *What are statistics for? Linking statistical learning to language acquisition in the wild*. Paper presented at the International Conference on Infant Studies, Chicago, IL.
- Saffran, J. R., & Griepentrog, G. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, 37(1), 74–85.
- Saffran, J. R., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2005). *wCross-species differences in the capacity to acquire language: Grammatical pattern learning by human infants and monkeys*. Manuscript submitted for publication.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- Saffran, J. R., Loman, M. M., & Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, 77, 15–23.
- Saffran, J. R., Reeck, K., Niehbur, A., & Wilson, D. P. (in press). Changing the tune: Absolute and relative pitch processing by adults and infants. *Developmental Science*, 7(7), 53–71.
- Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, 39(3), 484–494.
- Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, 4(2), 273–284.
- Sanes, D. H., & Walsh, E. J. (1998). The development of central auditory function. In E. W. Rubel, R. R. Fay, & A. N. Popper (Eds.), *Development of the auditory system* (pp. 271–314). New York: Springer Verlag.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69(2), 105–134.
- Schafer, G., & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development*, 69(2), 309–320.
- Schneider, B. A., Bull, D., & Trehub, S. E. (1988). Binaural unmasking in infants. *Journal of the Acoustical Society of America*, 83, 1124–1132.
- Schneider, B. A., Morriongiello, 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.
- Schneider, B. A., Trehub, S. E., Morriongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America*, 86, 1733–1742.
- Sebastián-Gallés, N., & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance*, 28(4), 974–989.
- Sebastián-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, 31, 18–32.
- Seidenberg, M., & Elman, J. L. (1999). Do infants learn grammar with algebra or statistics? *Science*, 284, 433.
- Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2003). Are there limits to statistical learning? *Science*, 300, 53–54.
- Shady, M. E., & Gerken, L. A. (1999). Grammatical and caregiver cues in early sentence comprehension. *Journal of Child Language*, 26, 1–13.
- Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, 41(4), 874–886.
- Shahidullah, S., & Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Human Development*, 36(1), 13–26.
- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, 25(1), 169–201.
- Shi, R., & Werker, J. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, 12(1), 71–76.
- Shi, R., & Werker, J. F. (2003). The basis of preference for lexical words in 6-month-old infants. *Developmental Science*, 6(5), 484–488.
- Shi, R., Werker, J. F., & Cutler, A. (2003). Function words in early speech perception. In *Proceedings of the 15th International Conference of Phonetic Sciences* (pp. 3009–3012). Adelaide, Australia: Causal Productions.
- Shi, R., Werker, J. F., Cutler, A., & Cruickshank, M. (2004, April). *Facilitation effects of function words for word segmentation in infants*. Poster presented at the International Conference of Infant Studies, Chicago.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), B11–B21.
- Shvachkin, N. K. (1948). The development of phonemic speech perception in early childhood. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 91–127). New York: Holt, Rinehart and Winston.
- Singh, L., Bortfeld, H., & Morgan, J. (2002). Effects of variability on infant word recognition. In A. H. J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 608–619). Somerville, MA: Cascadilla Press.
- Singh, L., Morgan, J., & White, K. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, 51, 173–189.
- Sininger, Y., & Abdala, C. (1996). Auditory brainstem response thresholds of newborns based on ear canal levels. *Ear and Hearing*, 17, 395–401.
- 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, 1–22.

106 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- Sinnott, J. M. (1989). Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus*, *Macaca*) and humans. *Journal of the Acoustical Society of America*, *86*, 557–565.
- Sinnott, J. M., & Aslin, R. N. (1985). Frequency and intensity discrimination in human infants and adults. *Journal of the Acoustical Society of America*, *78*, 1986–1992.
- Sinnott, J. M., Pisoni, D. B., & Aslin, R. M. (1983). A comparison of pure tone auditory thresholds in human infants and adults. *Infant Behavior and Development*, *6*, 3–17.
- Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York: Holt, Rinehart and Winston.
- Smith, S. L., Gerhardt, K. J., Griffiths, S. K., & Huang, X. (2003). Intelligibility of sentences recorded from the uterus of a pregnant ewe and from the fetal inner ear. *Audiology and Neuro Otolology*, *8*, 347–353.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267.
- Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, *61*, 632–652.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381–382.
- Stellmack, M. A., Willihnganz, 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.
- Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1321–1337.
- Strange, W., & Jenkins, J. J. (1978). Role of linguistic experience in the perception of speech. In R. D. Walk & H. L. Pick (Eds.), *Perception and experience* (pp. 125–169). New York: Plenum Press.
- Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, *259*, 39–41.
- Studdert-Kennedy, M., & Shankweiler, D. P. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, *48*(2, Pt. 2), 579–595.
- Swingle, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, *50*, 86–132.
- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166.
- Swingle, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484.
- Swingle, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*(2), 73–108.
- Swoboda, P., Kass, J., Morse, P. A., & Leavitt, L. A. (1978). Memory factors in infant vowel discrimination of normal and at-risk infants. *Child Development*, *49*, 332–339.
- Swoboda, P. J., Morse, P. A., & Leavitt, L. A. (1976). Continuous vowel discrimination in normal and at risk infants. *Child Development*, *47*(2), 459–465.
- Tallal, P., Miller, S. L., Jenkins, W. M., & Merzenich, M. M. (1997). The role of temporal processing in the developmental language-based learning disorders: Research and clinical implications. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 49–66). Mahwah, NJ: Erlbaum.
- Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, *241*, 468–469.
- Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*, 83–93.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.
- Tarquinio, N., Zelazo, P. R., & Weiss, M. J. (1990). Recovery of neonatal head turning to decreased sound pressure level. *Developmental Psychology*, *26*, 752–758.
- Tees, R. C. (1976). Perceptual development in mammals. In G. Gottlieb (Ed.), *Studies on development of behavior and the nervous system* (pp. 281–326). New York: Academic Press.
- Teller, D. Y. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development*, *2*, 135–153.
- Tharpe, A. M., & Ashmead, D. H. (2001). A longitudinal investigation of infant auditory sensitivity. *American Journal of Audiology*, *10*, 104–112.
- Thiessen, E. D. (2004). *The role of distributional information in infants' use of phonemic contrast*. Unpublished doctoral dissertation, University of Wisconsin, Madison.
- 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*(4), 706–716.
- Thiessen, E. D., & Saffran, J. R. (2004). Infants' acquisition of stress-based word segmentation strategies. In A. Brugos, L. Micciulla, C. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*, *2*, 608–619.
- Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *Journal of Speech Language and Hearing Research*, *42*, 1061–1068.
- Thorpe, L. A., & Schneider, B. A. (1987, April). *Temporal integration in infant audition*. Paper presented at the Society for Research in Child Development, Baltimore, MD.
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175.
- Toro, J. M., & Trobalón, J. B. (2004). *Statistical computations over a speech stream in a rodent*. Manuscript submitted for publication.
- Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, *7*, 289–296.
- Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology*, *9*(1), 91–96.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, *47*, 466–472.
- Trehub, S. E. (2003). Musical predispositions in infancy: An update. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 3–20). Oxford, England: Oxford University Press.
- Trehub, S. E., Endman, M. W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology*, *49*, 300–313.
- Trehub, S. E., Schneider, B. A., & Edman, M. (1980). Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology*, *29*, 282–293.

- 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.
- Trehub, S. E., Schneider, B. A., Morreniello, 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., Thorpe, L. A., & Judge, P. (1991). Observational measures of auditory sensitivity in early infancy. *Developmental Psychology*, *27*, 40–49.
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech Perception in Infancy Predicts Language Development in the Second Year of Life: A longitudinal study. *Child Development*, *75*, 1067–1084.
- Tsushima, T. T. O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., et al. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities. *Proceedings of International Conference of Spoken Language Processing, Yokohama*, *94*, 1695–1698.
- Turk, A. E., Jusczyk, P. W., & Gerken, L. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech*, *38*(2), 143–158.
- Vargha-Khadem, F., & Corballis, M. C. (1979). Cerebral asymmetry in infants. *Brain and Language*, *8*(1), 1–9.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *Journal of the Acoustical Society of America*, *66*, 1380–1564.
- Viemeister, N. F. (1996). Auditory temporal integration: What is being accumulated? *Current Directions in Psychological Science*, *5*, 28–32.
- Viemeister, N. F., & Plack, C. J. (1993). Time analysis. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (Vol. 3, pp. 116–154). New York: Springer-Verlag.
- Viemeister, N. F., & Schlauch, R. S. (1992). Issues in infant psychoacoustics. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 191–210). Washington, DC: American Psychological Association.
- Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, *90*, 858–865.
- Vihman, M., Nakai, S., dePaolis, R., & Hallè, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, *50*(3), 336–353.
- Vitevitch, M. S., & Luce, P. A. (1998). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408.
- Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*, 47–62.
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, *13*(7), 994–1005.
- Vouloumanos, A., & Werker, J. F. (2002, April). *Infants' preference for speech: When does it emerge?* Poster presented at the 13th biennial International Conference of Infant Studies, Toronto, Ontario, Canada.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science*, *7*(3), 270–276.
- Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory and Cognition*, *24*(6), 744–755.
- Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology*, *3*(2), 55–75.
- Walton, G. E., & Bower, T. G. R. (1993). Amodal representation of speech in infants. *Infant Behavior and Development*, *16*, 233–243.
- Waters, R. A., & Wilson, W. A., Jr. (1976). Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. *Perception and Psychophysics*, *19*, 285–289.
- Weir, C. (1976). Auditory frequency sensitivity in the neonate: A signal detection analysis. *Journal of Experimental Child Psychology*, *21*, 219–225.
- Weir, C. (1979). Auditory frequency sensitivity of human newborns: Some data with improved acoustic and behavioral controls. *Perception and Psychophysics*, *26*, 287–294.
- Werker, J. F., Burns, T., & Moon, E. (in press). *Bilingual exposed newborns prefer to listen to both of their languages*. Manuscript in preparation.
- 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*(6), 1289–1309.
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234.
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1–30.
- Werker, J. F., Frost, P. E., & McGurk, H. (1992). La langue et les lèvres: Cross-language influences on bimodal speech perception. *Canadian Journal of Psychology*, *46*(4), 551–568.
- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349–355.
- Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*(5), 672–683.
- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, *37*(1), 35–44.
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, *43*(2), 230–246.
- Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*(2), 278–286.
- 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*(1), 49–63.
- Werker, J. F., Weikum, W. M., & Yoshida, K. A. (in press). Bilingual speech processing. In W. Li (Series Ed.) & E. Hoff & P. McCardle (Vol. Eds.), *Multilingual Matters*.
- Werner, L. A. (1992). Interpreting developmental psychoacoustics. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 47–88). Washington, DC: American Psychological Association.
- Werner, L. A. (1995). Observer-based approaches to human infant psychoacoustics. In G. M. Klump, R. J. Dooling, R. R. Fay, &

108 The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language

- W. C. Stebbins (Eds.), *Methods in comparative psychoacoustics* (pp. 135–146). Boston: Birkhauser Verlag.
- Werner, L. A. (1999). Forward masking among infant and adult listeners. *Journal of the Acoustical Society of America*, *105*, 2445–2453.
- Werner, L. A. (2003, March). *Development of backward masking in infants*. Paper presented at the American Auditory Society, Scottsdale, AZ.
- 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*, 2101–2111.
- Werner, L. A., Folsom, R. C., & Mancl, L. R. (1993). The relationship between auditory brainstem response and behavioral thresholds in normal hearing infants and adults. *Hearing Research*, *68*, 131–141.
- Werner, L. A., Folsom, R. C., & Mancl, L. R. (1994). The relationship between auditory brainstem response latencies and behavioral thresholds in normal hearing infants and adults. *Hearing Research*, *77*, 88–98.
- Werner, L. A., & Gillenwater, J. M. (1990). Pure-tone sensitivity of 2- to 5-week-old infants. *Infant Behavior and Development*, *13*, 355–375.
- Werner, L. A., & Mancl, L. R. (1993). Pure-tone thresholds of 1-month-old human infants. *Journal of the Acoustical Society of America*, *93*, 2367.
- 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.
- Whalen, D. H., & Liberman, A. M. (1987). Speech perception takes precedence over nonspeech perception. *Science*, *237*, 169–171.
- 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., 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*, 3297–3305.
- Wilmington, D., Gray, L., & Jahrsdorfer, R. (1994). Binaural processing after corrected congenital unilateral conductive hearing loss. *Hearing Research*, *74*, 99–114.
- Wong, D., Miyamoto, R. T., Pisoni, D. B., Sehgal, M., & Hutchins, G. (1999). PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech stimuli. *Hearing Research*, *132*, 34–42.
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*, 553–566.
- 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*, 129–130.
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehler, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, *102*, 1161–1171.
- Zangl, R., & Fernald, A. (2003, October). *Sensitivity to function morphemes in on-line sentence processing: Developmental changes from 18 to 36 months*. Paper presented at the biennial meeting of the Boston University Conference on Child Language Development, Boston, MA.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*(5058), 846–849.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. (1996). PET studies of phonetic processing of speech: Review, replication, and re-analysis. *Cerebral Cortex*, *6*(1), 21–30.