Chapter 2
Infant and Childhood Development: Intersections Between Development and Language Experience

Fuh-Cherng Jeng

Abstract This chapter provides an overview of the development and plasticity of the neural encoding of speech and non-speech stimuli at the subcortical level with an emphasis on the influence of an individual's language experience during infancy and childhood. Sections of this chapter are developed based upon a theoretical framework to embrace all possible sources and interactions that may have a significant effect on the development of the auditory system during the early stages of life. The discussion begins with the acoustic environment of a human fetus and the possible influence of prenatal listening experience on the development of the auditory system at the subcortical level. Next, the development during an infant's immediate postnatal days and first year of life are presented. Developmental trajectories and possible influences of linguistic experience on speech processing, as illustrated by the many aspects of neural encoding (e.g., tracking acuity, pitch strength, and the spectral and timing accuracy at the fundamental frequency, harmonics, and speech formants) are discussed. The presentation continues through childhood, which manifests itself as exponential growth of the developmental trajectory and adaptation of the auditory system. Neural encoding of the various aspects of human speech is described as it pertains to children situated in a quiet or noisy acoustic environment. The effects of acquiring more than one language, sequentially or simultaneously, are also discussed.

Keywords Auditory system · Frequency-following response · Fundamental frequency · Harmonics · Innate ability · Linguistic experience · Neural circuitry · Neonate · Neuroplasticity · Newborn · Pitch · Speech formant · Subcortical

F.-C. Jeng (✉)
Communication Sciences and Disorders, Ohio University, Athens, OH 45701, USA
e-mail: jeng@ohio.edu

2.1 Introduction

After a gestational period, human infants are delivered to the "outside" world. Immediately, infants encounter novel sensory inputs. Interactions between the infants and their native environments take place promptly in multiple forms (Hollich et al. 2000). Human infants, equipped with astonishing anatomical structures and functional capacities (Eggermont and Moore 2012) have the ability to interact with their environment through modalities such as vision, touch, smell, and, clearly, sound. The sensory inputs that are available in the infant's native environment, together with the infant's interactions with caregivers, facilitate brain development, which includes the perception and production of prosody, music (Trainor and Unrau 2012), gesture, and later on, the mastery of a language (Panneton and Newman 2012).

Immediately after birth infants are exposed to the rich acoustic stimuli that are available in their native environments. Development of the auditory system thus intersects with the infant's exposure to their acoustic environment during the early stages of life. However, the pathway to acquire language is a long process. While it takes approximately three years for most children to master the perception and production skills necessary for communicating in their native languages (Jusczyk 1997), some children may take a longer time to reach the same speech and language proficiencies than their peers (see Schochat, Rocha-Muniz, and Filippini, Chap. 9). Acquiring language can be challenging when infants and children are placed in adverse acoustic environments, such as background noise and reverberation (see Bidelman, Chap. 8). When a person encounters difficulty processing specific aspects of speech, individualized training and rehabilitative protocols may be considered (see Carcagno and Plack, Chap. 4). With or without difficulty processing human speech during the early stages of life, the brain continues to develop during adolescence and adulthood (see Krishnan and Gandour, Chap. 3) and starts to decline during the senior years (see Anderson, Chap. 11).

This chapter provides an overview of the development and plasticity of the neural encoding of speech and non-speech stimuli at the subcortical level with an emphasis on the influence of an individual's language experience during infancy and childhood. The chapter begins with a brief description of the measurements that are utilized to examine the various aspects of speech processing at the subcortical level. The following sections are then developed based upon a theoretical framework to embrace all possible sources and interactions that may have a significant effect on the development of the auditory system during the early stages of life. The discussion begins with the acoustic environment of a human fetus and the possible influence of prenatal listening experience on the development of the auditory system at the subcortical level. Next, the development during an infant's immediate postnatal days and first year of life are presented, together with the recent documentations of the FFR (frequency-following response) literature, including longitudinal follow-up of speech encoding during the first year of life and cross-sectional studies that have included infants with various age groups. Developmental

trajectories and possible influences of linguistic experience on speech processing as illustrated by the various aspects of neural encoding will be discussed (e.g., tracking acuity, pitch strength, and the spectral and timing accuracy at the fundamental frequency (F_0), harmonics, and speech formants). Of course, the importance of the stability and precision of neural firing at the subcortical level cannot be omitted. The presentation continues through childhood—additional exponential growth of the developmental trajectory and adaptation of the auditory system. Neural encoding of the various aspects of human speech is described as it pertains to children who are situated in a quiet or noisy acoustic environment. Effects of acquiring more than one human language, sequentially or simultaneously, are also discussed. Lastly, current issues concerning the possible influence of prenatal listening experience, lack of age-appropriate normative databases, and the urgency of developing functional and computational models suitable for the different age groups during infancy and childhood are discussed.

2.2 Frequency-Following Response

Interactions between the development of a human brain and the listener's linguistic experience take place both cortically and subcortically. While many studies have focused on the development and neuroplasticity of the neural structures at the cortical level (He et al. 2007; Näätänen et al. 2007; Brauer et al. 2011; Butler and Trainor 2013; Partanen et al. 2013), few have emphasized the neuroplasticity and functional organization of the neural circuitries at the subcortical level.

An electrophysiological measurement that provides insight into the subcortical pitch processing mechanisms of the auditory system is known as the frequency-following response (FFR): a term first used by Worden and Marsh (1968) to describe a response that is phase locked to the frequency components of the stimulus in cats. Phase locking is a phenomenon indicating that neural structures are firing action potentials that are synchronized with the waveform morphology of a stimulus. The FFR was later used to define an electrophysiological measure of the human auditory system's ability to track the frequencies of 500 and 1,000 Hz pure tones (Moushegian et al. 1973). Generally speaking, FFR is a collective response from a group of neurons that responds to the periodicity of the stimulus waveform. The F_0 of the acoustic stimulus presented to a listener, if encoded accurately, appears as periodic peaks on the recorded waveform. The inter-peak intervals of the stimulus and corresponding response waveform are a reflection of the neural phase-locking abilities of the auditory system. Investigations showing the neural phase-locking properties of the FFR reflect the *temporal* and *place theories*. The temporal and place theories are thought to contribute to the brain's ability to encode spectral information from speech stimuli. Because the FFR is a scalp-recorded auditory evoked potential and requires no active participation of the listener, it serves as a noninvasive and objective measurement of the capacity that human

subcortical neurons have to receive and track changes in the frequency content embedded in an acoustic stimulus (Skoe and Kraus 2010).

Through the use of complex sounds (e.g., speech), various aspects of neural encoding at the subcortical level can be examined. Complex sounds elicit equally complex responses, offering a rich set of ingredients the brain has to process. Due to this nature, the FFR to complex sounds contains enriched information that may be useful to help decipher the complex, and yet dynamic, neural networks inside the brain. It is likely that the neural elements involved in the processing of complex speech sounds are not limited to one area of the brain but instead are distributed across various subcortical neural structures. These neural structures, although dispersed among the nuclei, respond concertedly with the various features of human speech. Examination of such complex but integrated neural circuitry requires a method that allows the investigation of the neural elements that respond synchronously with the incoming signal. The FFR provides such a method and opens a door to examine the various and yet distinct biological processing of speech sounds at the subcortical level. Details about the methodology and procedures of recording an FFR in response to complex sounds can be found in Chap. 1 by Kraus, Anderson, and White-Schwoch.

Developmental trajectories of subcortical pitch representation, as reflected by the scalp-recorded FFR, undergo stages of progression across a human lifespan. During the immediate postnatal days, newborns are able to process tones (Gardi et al. 1979) and the changes in voice pitch (Jeng et al. 2011b). Through infancy, pitch representations at the subcortical level undergo a phase of rapid maturational changes (Jeng et al. 2010; Anderson et al. 2015). During childhood, subcortical pitch representation continues to improve and sometimes is accompanied by an overshoot in the response latency and amplitude: shorter response latency and larger response amplitude than those observed in adults (Skoe et al. 2015). This overshoot is followed by a gradual increase in the response latency and a gradual decrease in the response amplitude that continues through adolescence. Pitch representation at the subcortical level remains relatively stable during adulthood but starts to decline when aging-related changes come into effect (see Anderson, Chap. 11). The following sections emphasize the possible influence of prenatal listening experience on subcortical pitch processing and the influence of postnatal linguistic experience on the development of neural responses to speech during infancy and childhood.

2.3 Gestational Age: Possible Influence of Prenatal Listening Experience

Before an infant is born, its mother's womb allows transmission of low-frequency sounds to the fetus (Gerhardt and Abrams 2000). Approximately three weeks after conception, the inner ear and cochlea start to take shape, and they become fully

functional by five months of gestation (Frenz et al. 2001); the fetus is then equipped with the sensory and neural elements necessary to receive acoustic information from the outside world. Thus, during the gestational period, the fetus is able to receive frequency-specific information from acoustic signals that are generated either by its mother or from others in its native environment. Because the fetus is submerged in an abundance of amniotic fluid, transmission of the acoustic signals likely will propagate to the fetus through the bone conduction pathway of the auditory system.

When a person vocalizes, vibration of his or her vocal folds creates sound waves. These sound waves then travel through the air in all directions. When the sound waves hit the mother's abdomen, most of the sound gets reflected because of the impedance mismatch between the air and the mother's abdominal tissue (Griffiths et al. 1994). However, if the sound producer is the fetus's mother herself, the sound waves generated by the mother can travel directly through her body to the fetus. In this case, acoustic sounds initiated by the mother can be transmitted, although attenuated, to the fetus (Gerhardt and Abrams 2000). Playback of intrauterine acoustic recordings demonstrates that the sounds available to the fetus are rich but somewhat muffled (Abrams et al. 1998a).

When acoustic vibrations travel through the mother's womb, not all frequencies become attenuated by the same amount. In other words, the loss of acoustic energy of the incoming signal is frequency specific (Abrams et al. 1998a). The mother's womb functions as a low-pass filter that allows frequencies below 300 Hz to reach the head of the fetus with very little or no attenuation (Abrams et al. 1998b); that is, for acoustic energies below 300 Hz, the intrauterine sound pressure is nearly identical to that generated by the mother's vocal folds. For acoustic energies above 300 Hz, the intrauterine sound pressure decreases with a slope of about 5 dB per octave. A decrease of 5 dB means that the amplitude of the intrauterine sound pressure becomes a little bit larger than half of its original amplitude. Thus, recordings of intrauterine sounds demonstrated relatively high intelligibility of externally generated human speech and music (Querleu et al. 1989; Griffiths et al. 1994). Although muffled, intrauterine sounds are still enriched with frequency-specific information, particularly at the low frequencies.

Toward the end of the second trimester and during the third trimester of a pregnancy, the mother often experiences the sensation of the fetus' movement in response to environmental sounds—consistent with results of scientific experiments where pure tones with various frequencies were delivered to fetuses while they were still in their mother's womb (Gelman et al. 1982; Shahidullah and Hepper 1994). Many experiments began at about the middle of the second trimester of a pregnancy, and the researchers were looking at the fetus' bodily reactions to sounds through an ultrasound. As time advanced and hearing progressed, fetuses responded to a greater range of frequencies and at a lower sound intensity. In Shahidullah and Hepper's (1994) study, they also examined the fetuses' responses to human speech. By observing the fetuses' habituation responses through ultrasounds, they found that fetuses at 35 weeks of gestational age were able to differentiate the pre-recorded human speech /baba/ versus /bibi/. This finding indicated that human

fetuses at 35 weeks' gestation already possessed the ability to discriminate between different phonemes. These informational cues may be helpful while the fetus is beginning to listen to sounds that are universal to all human languages and music. Some frequency components may even contain information that is specific to the native language of the fetus' external acoustic environment.

FFR recordings obtained in postnatal infants also provide evidence supporting the possible influence of prenatal listening experience on the development of the neural circuitry at the subcortical level. Anderson and colleagues (2015) reported that amplitudes of the FFRs at the first formant (F_1) and high harmonics (HH) increased significantly with increasing age, whereas the FFR amplitudes at the F_0 were clearly discernible in young infants and did not increase significantly in older infants. One possible explanation is that the low frequency vibrations (e.g., acoustic energies at F_0) are readily available during the prenatal stage of life, whereas the high frequency vibrations (e.g., acoustic energies at F_1 and HH) are unavailable to the fetus. The acoustic energies at the high frequencies (e.g., F_1 and HH) will become available to the fetus after birth. Exposure to these high frequency sounds in the listening environment after birth may have an effect on the growth of the F_1 and HH amplitudes during infancy. Importantly, the fact that the F_0 encoding is evident during early infancy corroborates the idea that the prenatal listening experience, at least for the low frequency components, may play an important role in facilitating normal development of the neural circuitry at the subcortical level during the early stages of life. This is also consistent with physiological evidence that low-frequency hearing sensitivity develops prior to high-frequency sensitivity in avian (Rubel and Ryals 1983) and mammalian (Echteler et al. 1989) systems.

2.4 Infancy

Immediately after infants are born they can detect almost all phonetic distinctions found in speech (Eimas et al. 1971; Kuhl et al. 2006). Interestingly, newborns and young infants exhibit a similar pattern of sound perception regardless of the language environment into which they are born (Kuhl 2010). This evidence indicates that the perception of speech is strongly influenced by innate factors (i.e., the *biological capacity model*). It is important to understand that the specific language environment to which an infant is exposed also effects perception of speech sounds (i.e., the *linguistic experience model*). Exposure to a specific language environment during the early stages of life results in a reduction in the ability to perceive differences among speech sounds of other languages (Kuhl et al. 1992; Kuhl 2004). For example, Kuhl and colleagues (1992) analyzed 6-month-old American and Swedish infants' perception of both native-language and foreign-language vowel sounds. They reported that the ability to hear differences among many of the sounds not used in the infant's language was lost by six months of age. During this same 6-month time frame, the infants' developmental speech perception of native sounds showed substantial enhancement and continued to do so until 12 months of age. For example, American

infants showed significant improvements in the discrimination of the English /r/-/l/ contrast in comparison to age-matched Japanese infants (Kuhl et al. 2006). Additionally, both the Chinese-learning and English-learning infants showed improvement on affricate-fricative contrasts between 6 and 12 months of age (Tsao et al. 2006). These linguistic experiences during the early stages of life, along with innate factors, play an important role in speech and language development.

2.4.1 Theories and Evidence About the Early Acquisition of Speech and Language

Universal traits and language-specific experiences both influence the acquisition of pitch perception. For theories related to the "biological capacity" model, Jakobson (1968) introduced the *law of irreversible solidarity*. This theory proposed that early acquisition of sound could be explained by the frequency distribution of that sound among the world's languages. Acoustic features that were more basic and central to all human languages, such as intonation, voice pitch, and rhythm, would be acquired earlier than the other aspects of speech sounds. Dinnsen (1992) suggested that there might be a *universal hierarchical structure* with a limited set of ordered acoustic features that were applicable to the inventories of all languages. Each feature in the hierarchy had a default (or unmarked) value. Therefore, acquisition of any feature of a specific language would involve a process of replacing a default value with a language-specific value. Dinnsen's (1992) model indicated that the order in which an infant acquires a specific feature of a language would depend on the dominant and default values of that feature. That is, acoustic features ranked high in the hierarchy would be acquired earlier than features ranked low. Jakobson's "law of irreversible solidarity" and Dinnsen's "universal hierarchical structure" emphasized the innate ability of humans to acquire language and were consistent with the "biological capacity" model.

In contrast, there are a number of theories that emphasize the role of perceptual importance on language acquisition. Locke (1980) proposed three mechanisms for language acquisition: maintenance, learning, and loss. Once an infant starts to acquire an account of targeted features of a language, certain sounds will be solidified within the infant's inventory. Sounds or specific voice patterns not present in the infant's early inventory are then learned through interactions within the postnatal linguistic experience. The infant will abandon and lose the sounds or certain voice patterns not present in their targeted language system. According to Locke, the interaction of these three mechanisms results in the acquisition of the targeted language to which the infant is exposed.

Kuhl (1994) proposed a *native language magnet theory* and then a revised version in 2008. The expanded version divides an infant's language development into four phases. *Phase 1* indicates that the infant's initial state is universally the

same. At birth, infants perceive sounds by their natural auditory processing mechanisms. During this phase, the infants' abilities do not depend on linguistic environment. *Phase 2* shows that early exposure to a specific language may cause physical changes in neural structure and circuitry, which become committed to recognizing acoustic patterns of native languages. For example, acoustic features that occur frequently in the infant's native language will stimulate certain neural structures repeatedly and may result in changes of specific neural structures and circuitry. Social interactions may also play a role in this phase by increasing the infant's attention and its awareness to specific acoustic patterns, and thus may have facilitated the functional reorganization of the infant's brain. *Phase 3* indicates how early linguistic experience repeatedly alters the initial state of the infant's perception of speech (i.e., magnet effects of speech perception). By six months of age, the infant's perception of speech not only deviates from the innate boundaries but also follows the distribution properties of sounds specific to its native language. *Phase 4* takes place when the neural commitment becomes stable. As infants come in contact with language-specific sounds, some form of this information is stored in their memory. A good example of this phase can be seen in Swedish, American, and Japanese infants. Behavioral responses of the infants were measured and showed that they produced distinctive representations that mirrored the distribution properties of ambient speech input. Over time, such magnet effects functionally erase certain boundaries that are irrelevant to the infant's native language (Kuhl et al. 2008).

One amazing feature of the human brain is its ability to adapt to the features of surroundings. For example, the acoustic and linguistic features of the listener's native language have substantial influences on the development of his/her processing of human speech. When infants are just born, they are capable of detecting subtle differences in speech sounds. That is, newborns can differentiate essentially all features of human speech (Eimas et al. 1971; Carral et al. 2005). Throughout the early stages of a human life, the brain develops and adapts to acoustic signals found in its environment. Such linguistic experiences initiate anatomical and functional refinement of the neural circuitry of the human brain. Over time, neural pathways that respond to the specific features of a language will be enhanced (Kuhl et al. 2008). For example, neonates who are born and raised in a tonal language environment will have substantial exposure to the distinctive intonation patterns (i.e., pitch contours) that are important in their native languages. Thus, neural circuits may be fine tuned to best respond to the pitch contours of the infant's linguistic environment. For example, Mandarin Chinese is a tonal language that utilizes distinctive pitch contours to deliver the different meanings of the same words. In Mandarin Chinese, there are four lexical pitch contours: *Tone 1*, *Tone 2*, *Tone 3*, and *Tone 4*. *Tone 1* has a flat pitch contour that remains relatively stable over its production. *Tone 2* starts from a low pitch utterance and gradually rises to a higher pitch. *Tone 3* has a falling and rising pitch contour with a reflection point around the mid portion of the utterance. *Tone 4* begins with a relatively high pitch utterance and gradually descends to a lower pitch. Each of the four pitch contours can carry a different meaning of the same word. For example, when the Mandarin syllable /yi/

is pronounced in Tone 1, it means "壹 [one]"; with Tone 2, it means "姨 [aunt]"; with Tone 3, it means "椅 [chair]"; and with Tone 4, it means "易 [easy]."

Behavioral studies have shown that infants learning tonal languages respond to changes in voice pitch in a categorical manner (Yip 2002; Panneton and Newman 2012). It was observed that four-months-old Chinese infants could discriminate the four lexical pitch contours with accuracy, and their ability in differentiating the four pitch contours persisted through infancy (Mattock and Burnham 2006). However, when American and French infants were tested with low versus rising lexical pitch contours in the Thai language, infants between four and six months old were able to discriminate the two lexical pitch contours with accuracy, but their performance dropped in infants who were nine months of age (Mattock et al. 2008).

2.4.2 *Electrophysiological Measurements During Infancy*

Electrophysiological measurements do not require behavioral feedback from the infant, enabling researchers to draw conclusions from an infant during the early stages of life. By about 4 to 5 months of age, an infant's brain is already sensitive to language-specific acoustic patterns and contrasts (Friederici et al. 2007). Electrophysiological studies that recorded cortical responses in infants indicated that early language exposure facilitates functional reorganization of brain networks (Grossmann et al. 2007; Friedrich and Friederici 2010). Although recordings of the cortical responses are useful in helping us to understand how the brain processes speech sounds at the cortical level, they have one drawback: they are affected by the state of the subject (Hall 2006). Thus, infants are required to remain awake throughout recordings (Friederici et al. 2007). This can be difficult to accomplish, particularly in young infants who have an attention span that is fairly short. One way to counterbalance this drawback is to record responses from neural structures at the subcortical level.

Although there is an abundance of literature documenting the behavioral and cortical responses in infants, few have focused on the functional organization at the subcortical level. Thus, the focus of the remainder of this chapter will be on the pitch processing and functional organization at the subcortical level for human infants. Electrophysiological studies that record responses from the subcortical neural structures have a major advantage over those that record responses from the auditory cortex. When recording responses from neural structures at the subcortical level, the infant does not have to stay awake or alert during data collection. Instead, the infant is encouraged to rest and fall asleep during research because alertness of the infant does not affect the results. An additional advantage of the FFR to complex sounds is that the FFR provides a precise representation of the stimulus features (Skoe and Kraus 2010), whereas the cortical response is an abstract representation (Hall 2006).

The FFR literature for neonates was first documented in 1979. During that time, Gardi and colleagues (1979) recorded the FFRs to low-frequency tone bursts in

full-term, healthy neonates. The FFRs recorded in neonates shared common characteristics in terms of the response morphology with those recorded in adults. That is, the response waveform followed the periodicity of the stimulus waveform. Latency of the first peak in the FFR decreased with increasing frequency. This was consistent with the general theories of hearing, the biomechanical properties of the basilar membrane, and the anatomy of the auditory system. The amplitude and threshold values of neonatal FFRs in response to low-frequency tone bursts are also similar to those obtained in normal-hearing adults. All together, these findings indicate that the integrity of the neural elements, particularly for those that are sensitive to low frequencies at the subcortical level, can be assessed starting from the first day of life.

2.4.3 Encoding of the Fundamental Frequency

After a gap of more than 30 years, characteristics of neonatal FFRs in response to speech stimuli were investigated. Jeng and colleagues (2011b) utilized a monosyllabic Mandarin stimulus that mimicked the English vowel /i/ and elicited FFRs in American and Chinese neonates during their immediate postnatal days. The FFRs were visualized by plotting spectral energies of the recordings as a function of time (Fig. 2.1). Spectrograms of the stimulus (Fig. 2.1, left column) clearly showed energy at the F_0 and its harmonics. Spectrograms of the recordings taken from 12 American and 12 Chinese neonates (Fig. 2.1, *middle column*) showed FFR energy that followed the fundamental frequency of the stimulus. The FFR energy following the harmonics was not as apparent. The FFRs recorded from both groups of neonates exhibited clear energy that followed the periodicity, such as the pitch contours, of the speech stimuli. Importantly, the FFRs obtained from American and Chinese neonates resembled each other and showed little differentiation. This finding provides evidence for the "biological capacity model," indicating that the neonates are born with similar innate abilities of pitch encoding at the subcortical level.

While the influence of the infant's linguistic environment has been shown through behavioral studies (Mattock and Burnham 2006; Mattock et al. 2008), little is known about how the various pitch contours are processed at the subcortical level and the developmental course during the first year of life. The influence of a person's linguistic experience on the development of the auditory circuitry at the subcortical level can be examined by recording FFRs in both neonates and adults who are born and raised in two different language environments. Cross-linguistic comparisons with the additional data that were obtained from 12 American and 12 Chinese adults (Fig. 2.1, *right column*) revealed the influence of a person's linguistic experience on the spectral encoding of a speech stimulus at the subcortical level. A study design including the four groups of participants (American neonates, American adults, Chinese neonates, and Chinese adults) allowed examination of the influence across and between maturity (neonate versus adult) and the listeners'
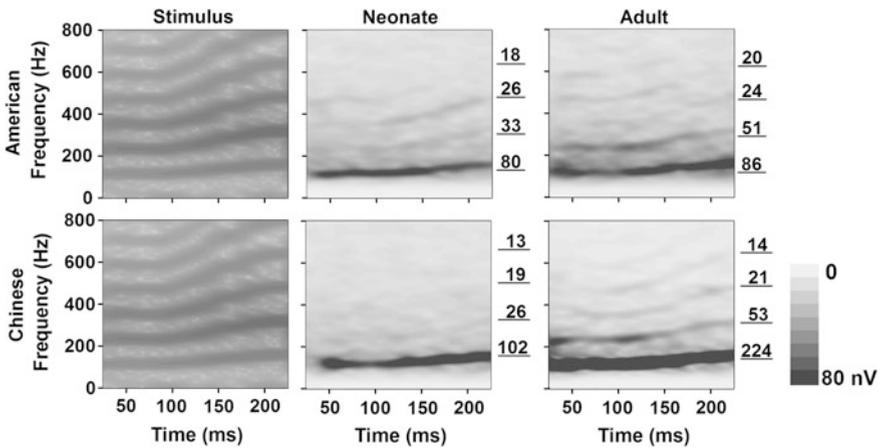
Fig. 2.1 Spectrograms of the stimulus (*left column*) and grand-averaged spectrograms obtained from 12 American neonates, 12 Chinese neonates (*middle column*), 12 American adults, and 12 Chinese adults (*right column*). The stimulus is a pre-recorded, monosyllabic, Mandarin speech stimulus that mimics the English vowel /i/ with a rising pitch. A *gray gradient scale* on the right of the spectrograms indicates the spectral amplitudes (nV) for the recordings obtained from neonate and adult participants. Underlined *numeric symbols* on the right of each spectrogram represent the means of the spectral amplitudes (nV) at the fundamental frequency, second, third, and fourth harmonics of the grand-averaged spectrograms. Spectral amplitudes of the harmonics were determined by finding the spectral peaks closest to those of the stimulus. The spectrograms of the stimulus are plotted on a normalized scale. All spectrograms were obtained using a Hanning window of 50 ms in length, overlap of 47.5 ms in length, and a frequency step of 1 Hz. (Reproduced with permission of publisher from Jeng et al. 2011b. © Ear and Hearing)

language experience (American versus Chinese) factors. The FFR data demonstrated a significant difference for the language factor, but not for the age factor nor for the interaction between the two factors. Furthermore, the FFRs obtained from the Chinese adults were significantly larger than Chinese neonates, whereas the FFRs obtained from American neonates and American adults were not significantly different from each other. These findings, together with the fact that American neonates and American adults had comparable FFRs, provide evidence supporting the "linguistic experience model."

Characteristics and maturational trends of the FFRs recorded in infants demonstrate the feasibility of studying the structural and functional reorganization of neural circuitry at the subcortical level during the first year of life. Jeng and colleagues (2010) recruited nine American infants, ranging from 1 to 11 months old. All infants were born and raised in native English-speaking households. The FFRs recorded from these infants showed discernible energy at the F_0 that followed the pitch contour of a speech stimulus. Four objective measures were applied to quantify the various aspects of pitch processing: *frequency error*, *slope error*, *tracking accuracy*, and *pitch strength*. Results obtained from the American infants were compared to those obtained in a group of American adults who were native

speakers of English. The four objective measurements were all focused on different aspects of pitch processing in the brainstem, and yet similar maturational trends were observed across all four measurements. Specifically, pitch-tracking acuity and phase-locking magnitude in infants appeared very similar to those in the adult population. This finding indicates that the neural circuitry needed to respond to the speech stimulus and formulate a discernible FFR morphology is readily available in infants.

Longitudinal follow-ups with each participant would potentially reveal a maturational trend for each individual. A prospective, longitudinal study with regular follow-ups was first made in a case study. Through a special opportunity, among the nine infants that were recruited in a previous study (Jeng et al. 2010), one infant was brought back for FFR evaluations at 1, 3, 5, 7, and 10 months of age. Spectrograms of the recordings obtained from this infant (Fig. 2.2) are arranged according to the age of the infant. As expected, FFRs recorded from individual listeners were not as robust as the grand-mean averages across all participants. Specifically, recordings obtained from individual listeners showed a relatively lower signal-to-noise ratio in the spectrogram, which resulted in occasional disruptions of the FFR in response to the F_0 contour of the stimulus (e.g., recordings obtained when this infant was 1 and 3 months old). This infant showed a weak response at 1 month old, but her responses became more visible at 3, 5, 7, and 10 months of age. This improvement suggested maturation in FFRs for infants, who do not show robust responses when they are young (e.g., 1 month old) but have developed strong pitch representation in the early stages of life (e.g., 3 months old). Objective indices regarding the tracking acuity and response magnitude of the FFRs obtained from this infant (Fig. 2.3) demonstrate a developmental trend of pitch encoding during the first year of life.

A cross-sectional study was performed and corroborated the idea of early maturation of neural encoding of F_0. Anderson and colleagues (2015) recorded FFRs to a synthesized /da/ syllable in 28 American infants 3–10 months of age. Results demonstrated that the F_0 amplitude in the FFR to the speech stimulus remained
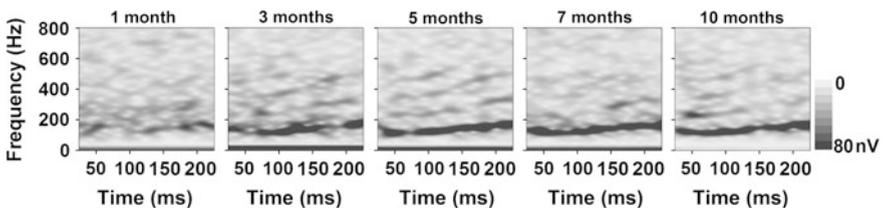
Fig. 2.2 Longitudinal follow-ups of an American infant revealed a maturational trend during the first year of life. This infant did not have a clearly identifiable FFR at 1 month old but showed a clearly identifiable FFR at 3 months old. After 3 months of age, this infant's FFR remained relatively stable. A gradient scale on the right indicates the spectral amplitudes (nV). All spectrograms were derived using a Hanning window of 50 ms in length, overlap of 47.5 ms in length, and a frequency resolution of 1 Hz. (Reproduced with permission of publisher from Jeng et al. 2010. © Perceptual and Motor Skills)
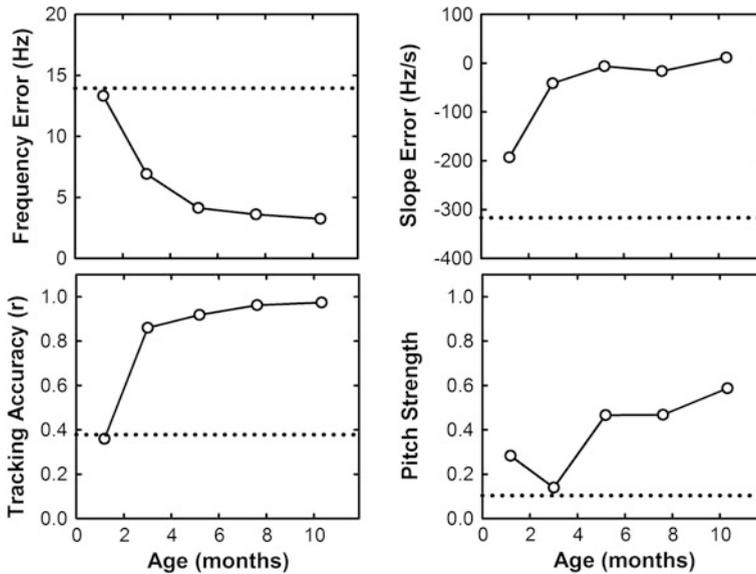
Fig. 2.3 Maturational trends of the FFRs obtained in an American infant at 1, 3, 5, 7, and 10 months of age. Frequency error, slope error, tracking accuracy, and pitch strength were computed from recordings obtained in this infant (see Fig. 2.2). Tracking acuity and phase-locking magnitude of the auditory system increased as this infant advanced in age. A *horizontal dotted line* within each panel indicates the mean of the control recordings, where the sound tube was occluded and moved away from the infant participant. (Reproduced from Jeng et al. 2010, with permission of publisher. © Perceptual and Motor Skills)

relatively stable for infants between 3 and 10 months of age. The growth of F_0 amplitude in the FFR occurs primarily during the early stages of life. This cross-sectional study demonstrated similar findings to those obtained in a longitudinal follow-up study of an infant. These findings signify that the development of neural circuitries responsible for encoding the F_0 information of a speech stimulus takes place early in time, likely sometime between 1 and 3 months old or even earlier.

2.4.4 Encoding Harmonics and Speech Formants

Frequency resolution is immature at birth for both animals and humans. Although the cochlea is fully functional at birth, neural elements at the subcortical level do not mature until later (Rubel and Ryals 1983). Specifically, neural elements sensitive to low frequencies emerge and mature earlier than those sensitive to high frequencies. Romand and Ehret (1990) studied the electrophysiological mapping in mice and reported that after birth the first recordable neural responses from the

inferior colliculus are at low frequencies. As the mice matured, frequency responsiveness of the neurons extended into the high frequencies. This postnatal development and maturation of the neural elements is likely driven by listening activities after birth and may have implications for the development of the encoding of harmonics and speech formants for human infants.

Neural structures at the subcortical level play an important role in deciphering harmonic and formant information for infants. Harmonics and speech formants are at higher frequencies than the F_0, and neural phase locking is clearer and more robust at low frequencies than high frequencies. For example, frequencies beyond 5000 Hz are too fast for any neuron to follow. Thus, when examining the harmonic or formant responses in an FFR, people direct their attention to frequencies below 5000 Hz. Although the characteristics of the harmonics and speech formants in the FFR have been reported in normal-hearing adults (Aiken and Picton 2006, 2008), few studies have examined the characteristics and implications of the harmonics and formants in the FFR for infants.

To date, only one paper reports the characteristics and development of the harmonics and formants in the FFR for infants. Anderson and colleagues (2015) recorded FFRs in American infants 3–10 months old. They reported that the amplitude of the F_0 in the FFR remained relatively stable, while the amplitudes of the first formant and high harmonics in the FFR increased as age increased (Fig. 2.4). Furthermore, when these infants were divided into two groups, younger and older (the younger group: 3–5 months old; the older group: 6–10 months old), the older infants demonstrated larger harmonic and formant amplitudes than the younger infants. These results not only provide evidence supporting improved neural encoding of speech features with age, but also highlight the importance of auditory neurodevelopment at the subcortical level for human infants.
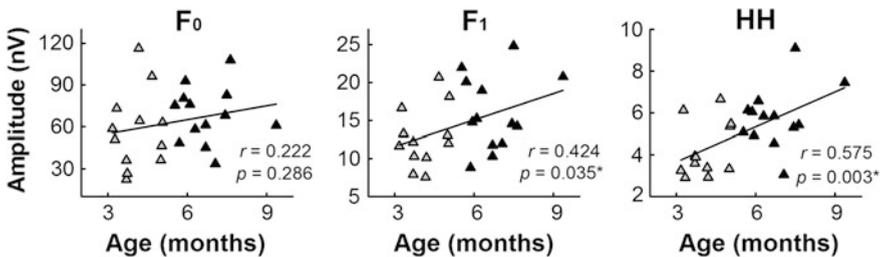
Fig. 2.4 Scatterplots of the F_0, F_1, and HH amplitudes of the FFRs recorded to a 40-ms /da/ syllable in 25 American infants, 3–10 months of age. No correlation was found between F_0 amplitudes of the FFRs and the age of the participating infants, but positive correlations were observed between the age of the infants and the F_1 and HH amplitudes of the FFRs. Younger infants (3–5 months old, *gray triangles*) have significantly smaller F_1 and HH amplitudes than older infants (6–10 months old, *black triangles*). *Solid lines* are the linear regressions of the FFR amplitudes as a function of age. *Asterisks* indicate $p < 0.05$; r, correlation coefficient; F_0, fundamental frequency; F_1, first speech formant; HH, high harmonics. (Reproduced from Anderson et al. 2015 with permission. © The Journal of the Acoustical Society of America)

Results of infant FFR studies are consistent with behavioral and other electrophysiological studies that have shown that frequency encoding of a human brain is immature at birth but improves during the first 6 months of age. For example, behavioral studies employing frequency discrimination tasks have reported that infants can detect frequency changes as small as 2–3% (Olsho et al. 1987). Importantly, infants at 3 months of age demonstrate a significantly larger frequency difference limen than those who are 6 months of age (Abdala and Folsom 1995). These findings indicate that the infant's ability to detect changes in frequency improves as they mature. Electrophysiological studies that record mismatch negativity responses from cortical neural structures confirmed the immature frequency resolution at birth followed by a significant improvement during the first few months.

2.4.5 Timing Aspects of the FFR

Timing aspects of neural encoding can be studied by examining the temporal waveform of the FFR. Gardi et al. (1979) utilized tone bursts (250, 500, and 1,000 Hz) and successfully recorded FFRs in neonates during their immediate postnatal days. Latency of the neonatal FFRs shared the characteristics of the physiological properties of the basilar membrane. Specifically, FFR latency decreased as a function of increasing frequencies of the tone bursts as in normal-hearing adults. Despite the similar characteristics shared by the neonates and adults, several differences were observed. For example, the neonatal FFRs elicited by 250 Hz and 500 Hz have longer latencies than those of normal-hearing adults. This is likely due to the fact that the neonate's brain remains under development, suggesting they require more input from their linguistic environment to further define the neural circuits of their auditory system.

In addition to examining the spectral components of the response, Anderson and colleagues (2015) investigated timing of onsets and offsets in response to a /da/ syllable in infants. It was discovered that the latency of the onset (A) peak, latency of the offset (O) peak, the inter-peak latency between the A and O peaks, and the onset slope from the V peak to the A peak were negatively correlated with the age of the infants (Fig. 2.5). Additionally, younger infants (3–5 months old) had longer peak latencies, shorter inter-peak A–O latency, and less abrupt onset of the V–A slope than older infants (6–10 months old). In other words, the younger the infant, the longer is the latency of the FFR and the less synchronous is the neural firing. These findings were consistent with the development of the human brain during infancy.

Although the cochlea is largely functional at birth, neural myelination and synaptic organization are still developing. For example, through the use of magnetic resonance imaging in infants, it has been reported that myelination of neural elements at the subcortical level takes place gradually during the first few months of life. For instance, the cochlear nucleus, superior olivary complex, and lateral
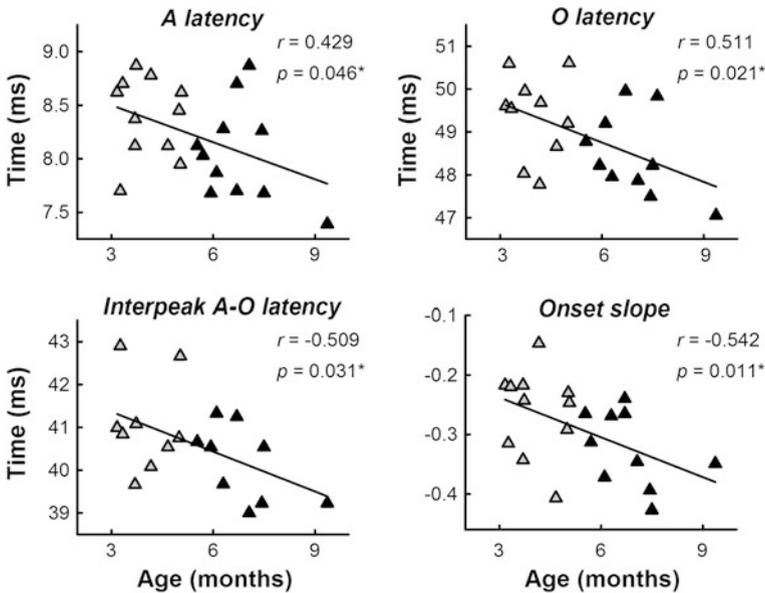
Fig. 2.5 Scatterplots demonstrating negative correlations between age and the latency of the *A peak* (onset), latency of the *O peak* (offset), interpeak latency between the *A* and *O peaks*, and the onset slope from the *peak V* to *peak A*. The FFR data were derived from recordings obtained in 25 American infants who were between 3 and 10 months old. Younger infants (3–5 months old, *gray triangles*) have significantly longer latencies than older infants (6–10 months old, *black triangles*). *Solid lines* are the linear regressions of the FFR measurements as a function of age. *Asterisks* indicate $p < 0.05$, r = correlation coefficient. (Reproduced from Anderson et al. 2015. © The Journal of Acoustical Society of America)

lemniscus show increased myelination density for up to 13 weeks of age, whereas the inferior colliculus shows improved myelination density for up to 39 weeks of age (Sano et al. 2007). The continued myelination of neural structures at the subcortical level may lead to the decreased latency of the FFRs recorded during the first year of life.

2.5 Childhood

Development and refinement of neural structures continue throughout childhood (Eggermont and Moore 2012). Specific speech characteristics and linguistic features that exist in the child's native language environment stimulate and enhance the functionalities of the neural elements and auditory circuitry on both the cortical (Huttenlocher and Dabholkar 1997) and subcortical (Song et al. 2008) levels. For the purpose of this chapter, discussion will be focused on subcortical changes. One major advantage of using a complex stimulus, such as the speech token /da/, is that

it allows a detailed examination of the timing and spectral properties of neural processing related to the abrupt onset of the consonant /d/, its transition to the vowel /a/, and the responses to the steady-state vowel portion of the stimulus. This approach provides enriched information about the neural processing of speech. As such, the FFR has been used as a neurophysiological marker that associates and predicts reading readiness, literacy capability, and academic performance for school-aged children and before they enter elementary school (White-Schwoch and Kraus 2013; White-Schwoch et al. 2015b). The FFR in response to complex stimuli also has been reported to be a viable method in assessing the impact of the child's linguistic environment on the functional and structural changes of the auditory neural circuitry for toddlers and school-aged children (Skoe et al. 2013; Krizman et al. 2015).

Precision of neural encoding at the subcortical level is correlated with a child's reading readiness and literacy (see Reetzke, Xie, and Chanderasekaran, Chap. 10). Additionally, the precision of neural encoding is decreased in children who have auditory processing disorders, dyslexia, or autistic spectrum disorders (see Schochat, Rocha-Muniz, and Filippini, Chap. 9). A more thorough discussion on these topics is available in the other chapters of the book (e.g., short-term learning and memory by Carcagno and Plack, Chap. 4; auditory experience and communication by White-Schwoch and Kraus, Chap. 6; and communicating in challenging environments by Bidelman, Chap. 8). The following sections of this chapter will emphasize normal development and the influences of linguistic experience on the subcortical neural encoding of speech during childhood.

2.5.1 Stability and Precision of Neural Firing

A reliable pattern of neural firing plays a pivotal role as a child continues to develop and adapt to the various, and sometimes adverse, listening environments they may encounter throughout childhood (White-Schwoch et al. 2015b). A consistent firing pattern among the distributed, but integrated, neural elements at the subcortical level is a physiological prerequisite for accurate encoding of the various linguistic and paralinguistic information embedded in human speech (Bidelman, Chap. 8). A child will be unable to process and perceive speech information accurately if the firing patterns of the involved neural elements fail to reflect the necessary and specific features of the speech sounds important in the child's native language.

Stability of neural firing can be quantified by measuring the *test-retest reliability* or *trial-by-trial variability* among the various sweeps of FFR recordings. While the test-retest reliability can be examined through recordings from the same participants within one test session, it can also be evaluated through analyzing the FFR recordings from the same participants but in different test sessions. An alternative approach is to examine the trial-by-trial variability by randomly selecting a fixed number of sweeps from a pool of all the recording sweeps obtained within a test session. The FFR literature from normally developing children reveals a reliable

and consistent firing pattern in a quiet listening environment (Russo et al. 2004; White-Schwoch et al. 2015a). It is worth mentioning that the various aspects of neural recording, as reflected by the different portions of FFR (e.g., the transient onset response versus the steady-state response) may have different levels of consistency in neural firing. This differentiation is important because speech perception is affected not only by how the brain processes a steady-state vowel but also by how a transient consonant is encoded by the brain. Russo and colleagues (2004) measured the test-retest reliability in eight normally developing children and found that FFR measurements derived from the sustained responses are more stable than those derived from the transient onset responses. Similar findings were discovered when comparing the trial-by-trial variability within a single FFR recording (White-Schwoch et al. 2015a). For example, Hornickel and Kraus (2013) reported a systematic relationship between the stability of the FFR and literacy skills, with reading-impaired children showing more variable responses to speech than their age-matched peers. Together, these results verify the notion that a consistent pattern of neural firing is a positive indicator regarding the maturity and readiness of the subcortical neural elements to receive and decipher the various specific features of human speech.

Efficiency and consistency of reliable firing patterns may be compromised when a child is listening in an adverse acoustic environment, such as in a reverberant room or in the presence of a substantial amount of background noise. White-Schwoch and colleagues (2015b) examined the FFRs recorded from a group of children 3–4 years old and reported that the precision and stability of the neural encoding of consonants in noise were strongly correlated with the children's phonological processing, reading readiness, and pre-literacy skills. Neural encoding of consonants in noise further predicted the children's performances on reading competence and a range of literacy tests when they grew older. When the data obtained from the group of children characterized by normal development were compared to the other children who had been diagnosed with a learning disability, the diagnostic group of children was found to have significantly poorer precision and stability of neural coding of consonants in noise when compared to the normal-developing children. Stability and precision of neural firing is a critical component for a child to receive and process the various features of speech stimuli consistently and efficiently.

2.5.2   Effects of Linguistic Experience

Neural structures of the auditory system are malleable such that a history of acoustic stimulation affects how neural circuits will respond to the subsequent incoming signals that may exist in a child's acoustic environment (Kral and Eggermont 2007; Shepard et al. 2013). Over time, changes of the neural structures in response to the acoustic signals of the listening environment will stay within the neural circuitry (Buonomano and Merzenich 1998; Kilgard and Merzenich 1998). This will

enhance the auditory processing of the linguistic and paralinguistic factors that are important in the child's native language. Neural elements at subcortical levels, similar to those at the cortical level, have a tendency and preference to fine tune to the acoustic and linguistic parameters that occur often in the child's language. The brain's top-down processing and the frequent stimulation of the linguistic and paralinguistic parameters on the neuronal structures are thought to facilitate the functional reorganization of auditory neural circuits at the subcortical level (Bidelman et al. 2011; Song et al. 2008).

A child's linguistic environment plays an important role in the development of the neural networks at the subcortical level (Krizman et al. 2012, 2015). In modern societies, it is nearly impossible to find a child who has no exposure to any language at all. Even tribes deep in a jungle communicate with native languages of their own. As a result, the influence of a child's linguistic experience on how the brain works is very difficult to isolate and examine. This restriction gives researchers no choice but to focus primarily on the differences between two languages. Through behavioral measurements and cortical responses to the different features of two languages, researchers have found that a child's experience and exposure to a specific language enhances auditory circuitries to the linguistic features found within that language (Zhang et al. 2005). At the same time, the child's ability to process the acoustic features that are specific and unique to another language will deteriorate if they are not present in the native language. With increased experience, the child's brain becomes fine-tuned to the linguistic and paralinguistic features that are specific and important in the child's native language. Admittedly, human communications are not limited to spoken languages. For example, people with profound hearing loss often communicate through sign languages (Sacks 1989). Development of the neural circuitries inside the brain for individuals who use sign languages involves interactions among visual and other sensory inputs (Kral and Eggermont 2007; Kral et al. 2013) and is beyond the scope of this chapter. For clarity, the term "language" is used in the remainder of this chapter to refer to spoken language.

In many instances, a child's linguistic environment contains more than one language. If a child communicates with two different languages, this child is called a bilingual. From the viewpoint of the FFR in response to complex sounds, studies have shown that children who are born and raised in a bilingual community exhibit stronger neural encoding to the spectral and timing components of speech (Fig. 2.6) (Krizman et al. 2012, 2015). They also have better consistency of neural firing than children who are born and raised in a mono-linguistic environment. Although, in most cases, one language is predominantly conversed throughout the child's daily life, the existence of the non-predominant language and the additions of bilingual factors create an enriched linguistic environment. An environment as such not only facilitates the acquisition of a second language for the child, but also further enhances the structural and functional reorganization of the auditory neural circuitries at the subcortical level during childhood (Kraus and White-Schwoch 2015; Krishnan and Gandour, Chap. 3).
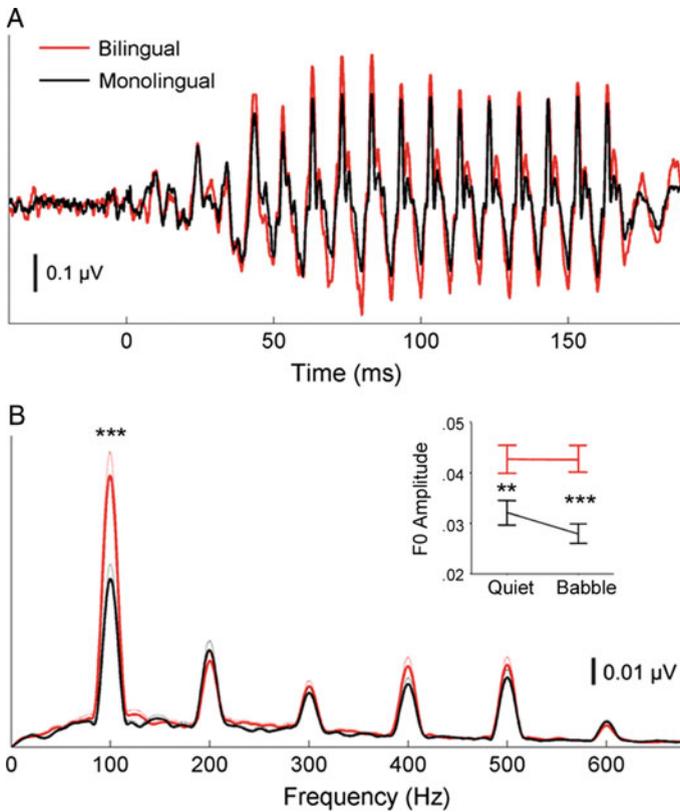
Fig. 2.6 Subcortical responses of bilinguals (*red*) and monolinguals (*black*) to the speech sound /da/ presented in multi-talker babble. (**A**) Bilinguals show a larger auditory brainstem response relative to monolinguals. (**B**) Amplitudes of the individual component frequencies in the steady-state (60–180 ms) region of the response to /da/ in multi-talker babble. *Thin lines* represent +1 standard error of the mean. *Inset in B* displays the mean amplitude (± 1 standard error) of the F_0 in quiet and in multi-talker babble for bilinguals and monolinguals. For monolinguals, there is a decrease in the amplitude of the F_0 (100 Hz) when the stimulus is presented in multi-talker babble relative to when it is presented in quiet. In contrast, bilinguals show virtually no change in F_0 amplitude between the two conditions. *Asterisks* represent significance levels: **$p < 0.005$, ***$p < 0.0001$. (Reproduced from Krizman et al. 2012 with permission. © Proceedings of the National Academy of Sciences of the USA)

Exposure to an additional language strengthens the auditory circuitry even further. Neural encoding and firing patterns are more resistant to the presence of background noise when a child has been exposed to an additional language (Krizman et al. 2012). For example, when speech sounds are presented in quiet, bilingual adolescents demonstrate larger FFR amplitudes and better consistency of neural firing. In the presence of background noise, such as multi-talker babble, the FFRs of bilinguals remain relatively stable. Monolinguals showed a decrement in

the accuracy and consistency of tracking the temporal and spectral aspects of human speech.

Children who acquire two spoken languages simultaneously from birth exhibit stronger neural encoding patterns that are phase locked to the spectral components of speech than those who acquire two languages sequentially (Krizman et al. 2015). Simultaneous bilinguals also demonstrate a greater consistency of neural firing than those who acquire two languages sequentially (Fig. 2.7). Additionally, the greater number of years of bilingual experience, the stronger neural encoding and the better trial-by-trial consistency of neural firing will be. Children with more years of experience communicating in two languages demonstrate stronger FFRs and more consistent firing patterns than those who have fewer years of bilingual experience, signifying the idea that enhanced spectral encoding and neural consistency will emerge with increasing years of experience communicating in a bilingual environment during childhood.
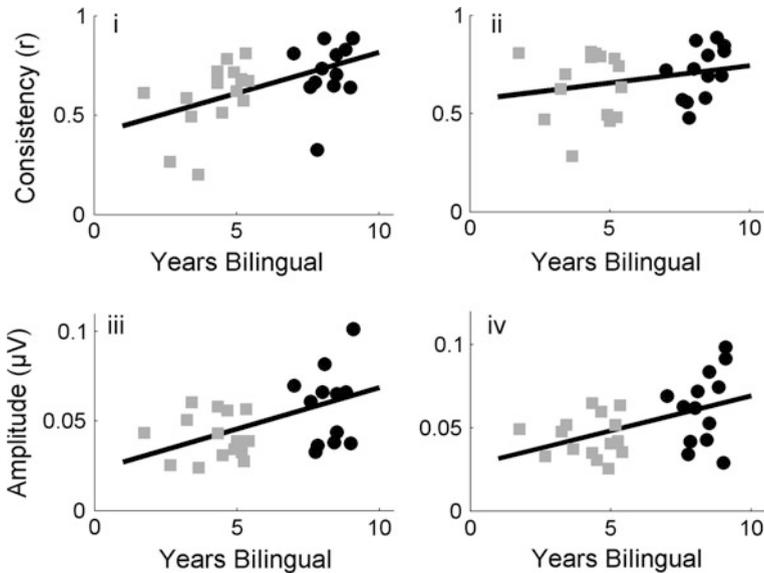
Fig. 2.7 Relationship between neural processing and years of bilingual experience. Response consistency for /ba/ (**i**) and /ga/ (**ii**) versus years of bilingual experience for the simultaneous (*black*) and sequential (*gray*) bilinguals plotted on the x-axis. The consistency to /ba/ relates with years of second language experience, while the consistency to /ga/ does not. F_0 encoding amplitude is shown for /ba/ (**iii**) and /ga/ (**iv**) versus years of bilingual experience for both groups. Both measures of F_0 encoding relate to the number of years of experience the child has speaking two languages. (Reproduced from Krizman et al. 2015 with permission of publisher. © Neuroscience Letters 2015)

2.6 Future Directions

Although the development of subcortical neural structures and the influence of linguistic experience during infancy and childhood have been extensively studied, a few issues remain.

The first issue is related to the listening experience of the fetus and its impact on the development of the auditory system. The human fetus becomes exposed to sounds of the linguistic environment while it is still in the mother's womb. It is hard to determine the influence of this linguistic exposure since there is no way to obtain an FFR in utero. Until recording FFRs in a living fetus becomes feasible, one can only derive conclusions indirectly. Currently, the earliest recording that can be made is during the immediate postnatal days. FFRs recorded in infants during their immediate postnatal days contain at least two components: one is the biological ability that comes with all newborns, and the other is the influence of the infant's listening experience when the infant is still in its mother's womb.

Although the FFR literature has provided partial and indirect evidence regarding the possible influence of prenatal listening experience on the development of auditory circuitry, a carefully designed study that includes the use of certain linguistic or musical training in pregnant women in a randomized-control study will be needed to properly address this issue. One possible approach is to develop an experimental protocol similar to that used in Partanen et al. (2013). A group of pregnant women could be recruited and asked to listen to a set of carefully designed sound material on a regular basis in order to familiarize the fetus with specific sound patterns (e.g., /tatata/ versus /tatota/). Electrophysiological responses could then be recorded in the neonates during their immediate postnatal days by using the specific sound patterns, and the responses could be compared to those elicited by other unrelated sound patterns.

The second issue is the lack of a normative database for each age group of normally developing infants and children. Completion and establishment of such a database for each age group is critical. Not only will this advance our understanding of the normal development of speech encoding at the subcortical level but also will allow the development of appropriate therapeutic and rehabilitative protocols for infants and children who are at risk of a specific disorder. Skoe and colleagues (2015) reported the development of subcortical auditory processing from 586 healthy participants across an extensive age range (ages 3 months to 72 years). This cross-sectional database is laudable and will be useful for future applications in the research and scientific realms. Additional databases supplementing the gaps in regards to the development of subcortical pitch processing during the immediate postnatal days and the first three months of life are warranted. Furthermore, a systematic large-scale multiple-site study, preferably prospective and longitudinal, will be needed to examine the characteristics and maturational trends of the FFR in normally developing infants and children across the various developmental stages of life. Infants and children who are born and raised in a non-tonal versus a tonal language environment and a mono-linguistic versus a multi-linguistic environment

should all be considered. Upon the completion and establishment of the normative database for FFRs, infants and children at risk for a specific disorder can be examined and their test results can be compared with those published in the age-appropriate normative database. Researchers and clinicians can further design treatment protocols with successful outcome measurements targeted to the normative database.

The third issue is related to the shortage of useable computational models that are capable of capturing the characteristics and growth trends of the FFRs for infants and children. Computational models are beneficial because they not only can help to understand how speech sounds are processed at the subcortical level, but they also can help predict outcomes of specific measurements of the brain. Ideal computational models should have solid foundations based on auditory anatomy and physiology. To initiate the process of developing computational models for FFRs, researchers have started testing some algorithms. For example, a computational model that utilizes an exponential curve-fitting formula has been successfully applied to normal-hearing adults (Jeng et al. 2011a), and an automatic procedure has been developed for neonates (Jeng et al. 2013). When performing tests on neonates, infants, and difficult-to-test populations, the amount of time needed to complete a recording is of great importance. Preliminary results have shown that the exponential curve-fitting model provides a good fit to the FFR trends with an increasing number of sweeps. Thus, the testing time can be shortened by employing an appropriate exponential model and applying a pre-determined stopping criterion to complete an FFR recording. However, further testing and finding other specific models that will work for the various age groups of participants will be needed to corroborate our understanding of speech encoding at the subcortical level and to predict outcomes in a simulated environment. A cadre of experts in auditory electrophysiology, computer modeling and simulation, pediatric neuroscience, language development, and related fields will need to collaborate to resolve this issue.

The last issue is associated with the inherently small amplitude of the scalp-recorded FFR. This issue becomes particularly challenging when attempting to record an FFR in a newborn nursery where environmental interferences can be substantial or when trying to record an FFR in an infant or a child who is not in a state of rest. Technologies that are designed to reduce the influence of environmental and other unwanted physiological noises and to enhance the robustness of the elicited response are needed. Algorithms necessary to promote the detection of the presence or absence of an FFR, along with automation of the necessary signal-processing procedures, will be needed to facilitate the visibility of an FFR. A real-time assessment of the progression of the data recording is preferred. Once the FFR algorithms and methodology have been further improved for detecting the presence of an FFR and its interface has become user friendly for researchers and clinicians in the FFR community and related fields, puzzles related to the normal development of the auditory system and related pathologies can be researched and resolved in a timely manner.

2.7 Summary

Since subcortical neural structures were reported to be malleable with auditory experience in the early twentieth century, a tremendous amount of new information and discoveries have been added to our understanding of the development of the human auditory system (Krishnan et al. 2005; Wong et al. 2007). Preliminary results obtained during the past 10 years have demonstrated possible impacts of language exposure during the early stages of life on the development of speech representation at the subcortical level. Future work in this area will benefit from collaborations among related disciplines and will promote a deeper understanding of the underpinning mechanisms involved in typical and atypical development of the auditory system during infancy and childhood.

Compliance with Ethics Requirements Fuh-Cherng Jeng declared that he had no conflict of interest.

References

Abdala, C., & Folsom, R. C. (1995). The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. *The Journal of the Acoustical Society of America, 98*, 921–930.

Abrams, R. M., Gerhardt, K. J., Griffiths, S. K., Huang, X., & Antonelli, P. J. (1998a). Intrauterine sounds in sheep. *Journal of Sound and Vibration, 216*, 539–542.

Abrams, R. M., Griffiths, S. K., Huang, X., Sain, J., et al. (1998b). Fetal music perception: The role of sound transmission. *Music Perception, 15*, 307–317.

Aiken, S. J., & Picton, T. W. (2006). Envelope following responses to natural vowels. *Audiology and Neurotology, 11*, 213–232.

Aiken, S. J., & Picton, T. W. (2008). Envelope spectral frequency–following responses to vowel sounds. *Hearing Research, 245*, 35–47.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2015). Development of subcortical speech representation in human infants. *The Journal of Acoustical Society of America, 137*, 3346–3355.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition, 77*, 1–10.

Brauer, J., Anwander, A., & Friederici, A. (2011). Neuroanatomical prerequisites for language functions in the maturing brain. *Cerebral Cortex, 21*, 459–466.

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience, 21*, 149–186.

Butler, B. E., & Trainor, L. J. (2013). Brief pitch-priming facilitates infants' discrimination of pitch-evoking noise: Evidence from event-related potentials. *Brain and Cognition, 83*, 271–278.

Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., et al. (2005). A kind of auditory 'primitive intelligence' already present at birth. *European Journal of Neuroscience, 21*, 3201–3204.

Dinnsen, D. (1992). Variation in developing and fully developed phonetics inventories. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 191–210). Timonium, MD: York Press.

Echteler, S. M., Arjmand, E., & Dallos, P. (1989). Developmental alterations in the frequency map of the mammalian cochlea. *Nature, 14*, 147–149.

Eggermont, J. J., & Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In L. Werner, R. R. Fay, & A. N. Popper (Eds.), *Human auditory development* (pp. 61–106). New York: Springer Science+Business Media.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303–306.

Frenz, D. A., McPhee, J. R., & Van De Water, T. R. (2001). Structural and functional development of the ear. In A. F. Jahn & J. Santos-Sacchi (Eds.), *Physiology of the ear* (pp. 191–214). San Diego: Singular Thomson Learning.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*, 1208–1211.

Friedrich, M., & Friederici, A. D. (2010). Maturing brain mechanisms and developing behavioral language skills. *Brain and Language, 114*, 66–71.

Gardi, J., Salamy, A., & Mendelson, T. (1979). Scalp-recorded frequency-following responses in neonates. *Audiology: Journal of Auditory Communication, 18*(6), 494–506.

Gelman, S. R., Wood, S., Spellacy, W. N., & Abrams, R. M. (1982). Fetal movements in response to sound stimulation. *American Journal of Obstetrics and Gynecology, 143*, 484–485.

Gerhardt, K. J., & Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *Journal of Perinatology, 20*, S20–S29.

Griffiths, S. K., Brown, W. S., Jr., Gerhardt, K. J., Abrams, R. M., & Morris, R. J. (1994). The perception of speech sounds recorded within the uterus of a pregnant sheep. *The Journal of Acoustical Society of America, 96*, 2055–2063.

Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: A neurobehavioral study. *Brain and Cognition, 64*, 30–41.

Hall, J. W., III. (2006). *New handbook of auditory evoked responses*. New York: Pearson.

He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience, 19*, 878–892.

Hollich, G., Hirsh-Pasek, K., Golinkoff, R., Brand, R. J., et al. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development* (Serial No. 262), *65*, 1–123.

Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: A biological marker of dyslexia. *The Journal of Neuroscience, 33*, 3500–3504.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*, 167–178.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague and Paris: Mouton.

Jeng, F.-C., Chung, H.-K., Lin, C.-D., Dickman, B. M., & Hu, J. (2011a). Exponential modeling of human frequency-following responses to voice pitch. *International Journal of Audiology, 50*, 582–593.

Jeng, F.-C., Hu, J., Dickman, B. M., Montgomery-Reagan, K., et al. (2011b). Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear and Hearing, 32*, 699–707.

Jeng, F.-C., Peris, K. S., Hu, J., & Lin, C.-D. (2013). Evaluation of an automated procedure for detecting frequency-following responses in American and Chinese neonates. *Perceptual and Motor Skills, 116*, 456–465.

Jeng, F.-C., Schnabel, E. A., Dickman, B. M., Hu, J., et al. (2010). Early maturation of frequency-following responses to voice pitch in infants with normal hearing. *Perceptual and Motor Skills, 111*, 765–784.

Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge (MA): MIT Press.

Kilgard, M. P., & Merzenich, M. M. (1998). Plasticity of temporal information processing in the primary auditory cortex. *Nature Neuroscience, 1*, 727–731.

Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews, 56*, 259–269.

Kral, A., Popper, A. N., & Fay, R. R. (2013). *Deafness*. New York: Springer Science+Business Media.

Kraus, N., & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: A cognitive-sensorimotor-reward framework. *Trends in Cognitive Sciences, 19*, 642–654.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25*, 161–168.

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the USA, 109*, 7877–7881.

Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience Letters, 585*, 48–53.

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology, 4*, 812–822.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843.

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67*, 713–727.

Kuhl, P., Conboy, B., Coffrey-Corina, S., Padden, D., et al. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*, 979–1000.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., et al. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21.

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.

Locke, J. (1980). The prediction of child speech errors: Implications for a theory of acquisition. In G. H. Heni-komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology. I. Production*. New York: Academic Press.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*, 241–265.

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*, 1367–1381.

Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and Clinical Neurophysiology, 35*, 665–667.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*, 2544–2590.

Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *The Journal of the Acoustical Society of America, 82*, 454–464.

Panneton, R., & Newman, R. (2012). Development of speech perception. In L. Werner, R. R. Fay, & A. N. Popper (Eds.), *Human auditory development* (pp. 197–222). New York: Springer Science+Business Media.

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., et al. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the USA, 110*, 15145–15150.

Querleu, D., Renard, X., Boutteville, C., & Crepin, G. (1989). Hearing by the human fetus? *Seminars in Perinatology, 13*, 409–420.

Romand, R., & Ehret, G. (1990). Development of tonotopy in the inferior colliculus. I. Electrophysiological mapping in house mice. *Developmental Brain Research, 54*, 221–234.

Rubel, E. W., & Ryals, B. M. (1983). Development of the place principle: Acoustic trauma. *Science, 4*, 512–514.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115*, 2021–2030.

Sacks, O. (1989). *Seeing voices: A journal into the world of the deaf*. New York: Vintage Books.

Sano, M., Kaga, K., Kuan, C.-C., Ino, K., & Mima, K. (2007). Early myelination patterns in the brainstem auditory nuclei and pathway: MRI evaluation study. *International Journal of Pediatric Otorhinolaryngology, 71*, 1105–1115.

Shahidullah, S., & Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Human Development, 36*, 13–26.

Shepard, K. N., Kilgard, M. P., & Liu, R. (2013). Experience-dependent plasticity and auditory cortex. In Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), *Neural correlates of auditory cognition* (pp. 293–327). New York: Springer Science+Business Media.

Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: A tutorial. *Ear and Hearing, 31*, 302–324.

Skoe, E., Krizman, J., Anderson, S., & Kraus, N. (2015). Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex, 25*, 1415–1426.

Skoe, E., Krizman, J., & Kraus, N. (2013). The impoverished brain: Disparities in maternal education affect the neural response to sound. *The Journal of Neuroscience, 33*, 17221–17231.

Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience, 20*, 1892–1902.

Trainor, L. J., & Unrau, A. (2012). Development of pitch and music perception. In L. Werner, R. R. Fay, & A. N. Popper (Eds.), *Human auditory development* (pp. 223–254). New York: Springer Science+Business Media.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America, 120*, 2285–2294.

White-Schwoch, T., Davies, E. C., Thompson, E. C., Woodruff Carr, K., et al. (2015a). Auditory-neurophysiological responses to speech during early childhood: Effects of background noise. *Hearing Research, 328*, 34–47.

White-Schwoch, T., & Kraus, N. (2013). Physiologic discrimination of stop consonants relates to phonological skills in pre-readers: A biomarker for subsequent reading ability? *Frontiers in Human Neuroscience, 7*, 1–9.

White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., et al. (2015b). Auditory processing in noise: A preschool biomarker for literacy. *PLoS Biology, 13*(7), e1002196. Doi:10.1371/journal.pbio.1002196

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.

Worden, F. G., & Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology, 25*, 42–52.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage, 26*, 703–720.

The Frequency-Following Response
A Window into Human Communication
Kraus, N.; Anderson, S.; White-Schwoch, T.; Fay, R.R.;
Popper, A.N. (Eds.)
2017, XVI, 294 p. 66 illus., 32 illus. in color., Hardcover
ISBN: 978-3-319-47942-2