Development/Plasticity/Repair

Developmental Plasticity in the Human Auditory Brainstem

Krista L. Johnson,[1,2] **Trent Nicol,**[1,2] **Steven G. Zecker,**[1,2] **and Nina Kraus**[1,2,3,4,5]

[1]Auditory Neuroscience Laboratory, [2]Department of Communication Sciences, [3]Institute for Neuroscience, and Departments of [4]Neurobiology and Physiology and [5]Otolaryngology, Northwestern University, Evanston, Illinois 60208

Development of the human auditory brainstem is thought to be primarily complete by the age of ∼2 years, such that subsequent sensory plasticity is confined primarily to the cortex. However, recent findings have revealed experience-dependent developmental plasticity in the mammalian auditory brainstem in an animal model. It is not known whether the human system demonstrates similar changes and whether experience with sounds composed of acoustic elements relevant to speech may alter brainstem response characteristics. We recorded brainstem responses evoked by both click and speech syllables in children between the ages of 3 and 12 years. Here, we report a neural response discrepancy in brainstem encoding of these two sounds, observed in 3- to 4-year-old children but not in school-age children. Whereas all children exhibited identical neural activity to a click, 3- to 4-year-old children displayed delayed and less synchronous onset and sustained neural response activity when elicited by speech compared with 5- to 12-year-olds. These results suggest that the human auditory system exhibits developmental plasticity, in both frequency and time domains, for sounds that are composed of acoustic elements relevant to speech. The findings are interpreted within the contexts of stimulus-related differences and experience-dependent plasticity.

Key words: auditory processing; ABR; speech encoding; plasticity; brainstem; development

Introduction

The auditory brainstem is a series of spatially separate nuclei that receive auditory input from the acoustic nerve and process this signal as it enters the neocortex. An important structure within this chain is the midbrain inferior colliculus (IC), because it acts as the primary relay center between ascending projections from the lower brainstem nuclei and ascending projections to the thalamus. Because of its unique makeup of converging ascending and corticofugal projections (Irvine, 1992; Zhang and Suga, 2005), animal studies have identified the IC as a site of both activity- and experience-dependent developmental plasticity in both vertebrate and, most recently, mammalian brains (Brainard and Knudsen, 1993; Knudsen, 1998; Zheng and Knudsen, 1999; De-Bello et al., 2001; Yu et al., 2007). However, similar studies on developmental plasticity in the human auditory brainstem have been limited primarily because of the obvious limitations in the use of parallel technique and experimental design. To gain understanding of plasticity in the human system, here we use speech-evoked electrophysiologic activity from the auditory brainstem to investigate whether the human auditory brainstem undergoes developmental changes that may be experience dependent.

The auditory brainstem response (ABR) is a noninvasive mea-

sure of far-field representation of stimulus-locked, synchronous electrical events. In response to an acoustic signal, a series of potential fluctuations measured at the scalp provide information about the functional integrity of brainstem nuclei along the ascending auditory pathway, making it a widely used clinical measure of auditory function (Despland and Galambos, 1980; Jacobson, 1985; Hood, 1998). Temporal precision is such that differences on the order of fractions of milliseconds are diagnostically significant. In response to click and tonal stimuli, a waveform will emerge with five major peaks of activity. Because the ABR is an aggregate neural response, it is difficult to identify with certainty the neural correlates of each of the five peaks. However, it is widely accepted that the first peak is generated by the auditory nerve and that the culmination of the synchronous activity resulting in the fifth peak is generated primarily within the midbrain IC (Jacobson, 1985).

There are well documented changes that occur with development over the I–V onset complex of the ABR (for review, see Salamy, 1984). Specifically, during the first 2 years of life, peak latencies become progressively earlier and peak amplitude increases. In terms of structural brainstem development, these ABR latency changes are thought to occur because of the rapid increase in axonal myelin density in the cochlear nerve and brainstem pathways (Moore and Linthicum, 2007). Interpeak timing differences (e.g., earlier peaks mature earlier) support a peripheral-to-central developmental trajectory. By the age of 2 years, the click and tone-evoked ABR waveform is fully mature and resembles that of an adult. Consequently, it is generally thought that functional development of the human auditory brainstem is complete at this time and, by inference, that the brainstem response to sound in general is mature by the age of 2 years.

More recently, the ABR has been used in humans to assess the

central processes that underlie complex signal patterns, such as those in speech and music. Although the waveform that emerges in response to a speech signal is similar to the waveform elicited by a click or a tone, it is more complex and actually mimics the acoustic properties of a speech syllable with remarkable fidelity (Kraus and Nicol, 2005; Johnson et al., 2007). A number of studies have demonstrated that click and speech stimuli impose different encoding demands on the brainstem. Specifically, a subset of children with learning and literacy disorders show abnormal neural encoding of speech in the presence of a normal click-evoked brainstem response (Cunningham et al., 2001; King et al., 2002; Wible et al., 2004, 2005; Banai et al., 2005; Johnson et al., 2005, 2007; Russo et al., 2005; Song et al., 2006). Of particular interest are the peak latency differences that occur within the first 10 ms of the response, the portion of the neural response thought to be most congruent across stimuli and generated within the IC. This evidence underscores an important neural encoding discrepancy between click and speech stimuli, despite similar generation sites. Furthermore, another component of the speech-evoked brainstem activity, the frequency-following response (FFR), has been recorded to speech in adults (Galbraith et al., 1995, 1997; Krishnan, 1999, 2002), but the developmental time course is unknown. The FFR reflects encoding of the fundamental frequency and harmonic structure of complex stimuli and also has midbrain origins (Galbraith, 1994).

Recent evidence from Yu et al. (2007) suggests developmental experience-dependent plasticity in the IC of mice, raising the possibility that similar mechanisms may exist in humans. One of the key differences between a click and speech stimulus is the environmental relevance of and exposure to the sounds. Moreover, it is known that higher-level cognitive activities such as language and music experience shape subcortical sensory infrastructure, notably the auditory brainstem response (Krishnan et al., 2005; Xu et al., 2006; Musacchia et al., 2007; Wong et al., 2007). To investigate the development of brainstem activity to speech and its relationship to the well known click developmental trajectory, click and speech-evoked ABRs were evaluated in children between the ages of 3 and 12 years. Specifically, we asked whether the neural response to a sound that is composed of acoustic elements relevant to speech has a different maturational time course than sounds known to produce a mature response well before the age of 3 years.

Materials and Methods

Institutional review board approval for this study was obtained from Northwestern University. Parental consent and the child's assent were obtained for all evaluation procedures, and children were paid for their participation in the study.

Participants. A total of 104 subjects between the ages of 3–5 and 8–12 years participated in this study. For the 3- to 5-year-old subjects, each year was treated as a separate age group, and each subject was represented in only one age group (i.e., no longitudinal data were collected). Subjects were aged 3 ($n = 22$), 4 ($n = 22$), and 5 ($n = 16$) years. The 8- to 12-year-old children were grouped together ($n = 44$), because no age differences have been found in this population (Russo et al., 2004). No children had histories of hearing loss, chronic ear infections, neurological disorders, or learning/attention problems. On the day of testing, subjects exhibited normal bilateral hearing (pure tone thresholds <20 dB HL for octaves between 500 and 4000 Hz and/or passed a screening pass/fail OAE). Moreover, all children had click-evoked brainstem response within normal limits [80 dB sound pressure level (SPL)]. Additional exclusionary criteria included learning and/or attention problems among immediate family members (parents and siblings). Subjects were considered normal learning based on information provided by a parent or guardian as gathered through a variety of reports.
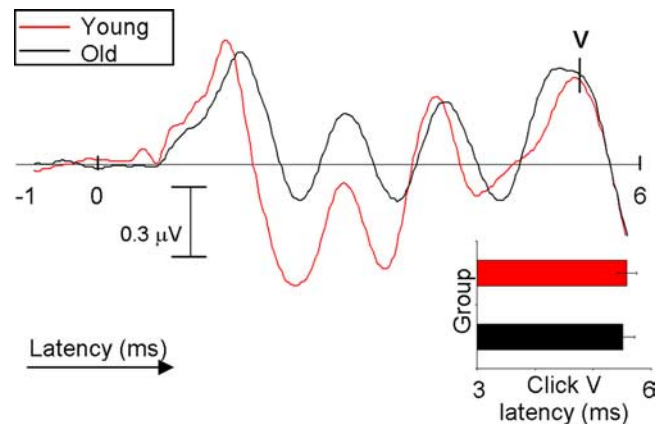

Figure 1. Representative click-evoked waveforms from a 3-year-old (red) and a 12-year-old (black) subject. Bar graph represents mean ± 1 SD for the young and old groups. There are no significant differences in click-evoked peak V latency.

Stimuli and recording. Brainstem responses were collected to both a click stimulus and a speech sound (/da/) according to widely used procedures as described in detail by Hood (1998) and Jacobson (1985). A Biologic Navigator Pro (Bio-logic, Mundelein, IL) was used to collect all physiological data. The Navigator's BioMAP (Biological Marker of Auditory Processing) module was used to collect the /da/-evoked responses. BioMAP uses a Klatt-synthesized (Klatt, 1980) 40 ms /da/ stimulus consisting of five formants with an onset burst frication during the first 10 ms at F_3, F_4, and F_5, and a fundamental frequency range of 105–121 Hz. The brainstem responses to speech and clicks were elicited by alternating polarities.

The test stimuli were presented to the right ear through Etymotic ER-3 earphones (Etymotic Research, Elk Grove Village, IL) at an intensity of 80 dB SPL. The left ear was unoccluded. To ensure subject cooperation and to promote stillness, all subjects watched videotaped programs, such as movies or cartoons, with the sound presented at a low level (<40 dB SPL). They were instructed to attend to the video rather than to the stimulus.

Recordings were made with silver–silver chloride electrodes (impedance <5 kΩ). Responses were differentially recorded from Cz-to-ipsilateral earlobe, with forehead as ground. Three blocks of 2000 artifact-free responses were collected at a rate of 13.3/s (click) and 10.9/s (/da/). For the click, a 10.66 ms recording window was used (including a 0.8 ms prestimulus period), and responses were on-line filtered from 100 to 1500 Hz. For the /da/, a 74.67 ms recording window (including a 15 ms prestimulus period) was used. Responses were sampled at 6856 Hz and bandpass filtered on-line from 100 to 2000 Hz. For both stimuli, sweeps with activity exceeding ±23.8 μV were rejected from the average. The three blocks were averaged after each recording session to yield a final waveform. Responses recorded in this study are thought to arise primarily from the auditory brainstem because of the filter characteristics and stimulation rates used. However, when recording such responses from the scalp, it is impossible to delineate the exact neural origin, such that cortical contributions cannot be ruled out.

Analysis. For the click-evoked response, peak latency and amplitude for wave V were identified for each subject. As is typically used in clinical evaluations, peak V was identified as the final data point on the waveform before the negative slope that follows the wave (Hall, 1992). The brainstem response to the speech sound /da/ has been described in detail in previous reports (Cunningham et al., 2001; King et al., 2002; Russo et al., 2004, 2005; Wible et al., 2004, 2005; Banai et al., 2005; Johnson et al., 2005, 2007) and is very reliable between and within subjects. Transient response measures include peak latency and amplitude measures. For each subject, peak latency and amplitude were determined for the brainstem onset (peaks V and A), offset (peak O), and the frequency-following peaks (D, E, F). A peak was deemed reliable if it was present in >85% of the total subject population. Peak C was deemed to be unreliable, because 18% of the entire subject population did not indicate a clear peak. Therefore, statistics were not performed on peak C. The VA onset complex was

further analyzed by computing slope and inter-peak latency, which are measures of neural synchrony for onset responders.

The sustained FFR to /da/ encodes the ongoing harmonic information within the speech syllable. This region was analyzed using two measures. First, the root-mean-square (RMS) amplitude was calculated over the time range of 21.9–40.6 ms. This was used to quantify the overall magnitude of the sustained activity, providing a measure of an individual's neural population response. Second, fast Fourier transform (FFT) analysis of the response was performed over the same time period to evaluate the spectral composition of the response. Average response magnitudes were calculated for 100-Hz-wide bins surrounding the frequency of the stimulus F_0 and the subsequent nine harmonics.

Descriptions of all analyses are provided throughout Results and are not repeated here.

Results

The latency of peak V of the click-evoked brainstem response was not significantly different among the four age groups (3-, 4-, 5-, and 8- to 12-year-old children) (ANOVA, $F = 2.057$, $p = 0.111$). This finding supports previous literature stating that peak V latency is mature and adult-like by the age of 2 years (Salamy, 1984). However, when the same analysis was performed on the onset portion of the speech-evoked response (peaks V and A), significant between-group differences were found (ANOVA, $F = 6.928$, $p < 0.001$; $F = 8.585$, $p < 0.001$, respectively). Least significant difference *post hoc* analyses showed no significant peak V or A latency differences between the 3- and 4-year-old groups or between 5-year-old and 8- to 12-year-old groups ($p > 0.05$), but both the 3- and 4-year-old groups had significantly delayed latencies for waves V and A compared with both the 5-year-old and 8- to 12-year-old groups (for corresponding p values, see Table 1). These findings suggest that brainstem neurons not only react differently to the onset of sound depending on whether it is a click or speech but also that 3- and 4-year-old children are representing the onset of a speech sound differently than 5- to 12-year-old subjects. Because this strong bimodal distribution between ages emerged in the speech-evoked response, age groups were collapsed to form two groups for the remainder of the analyses (the 3- and 4-year-old subjects were combined to create a "young" group, and the 5- to 12-year-old subjects were combined to create an "old" group). The latency of peak V for the click responses was still not significant between the young and old groups. Figure 1 shows click-evoked waveforms from a representative 3-year-old child and a representative 12-year-old child, as well as a bar graph illustrating the young versus old click peak V mean latencies.

Transient response measures

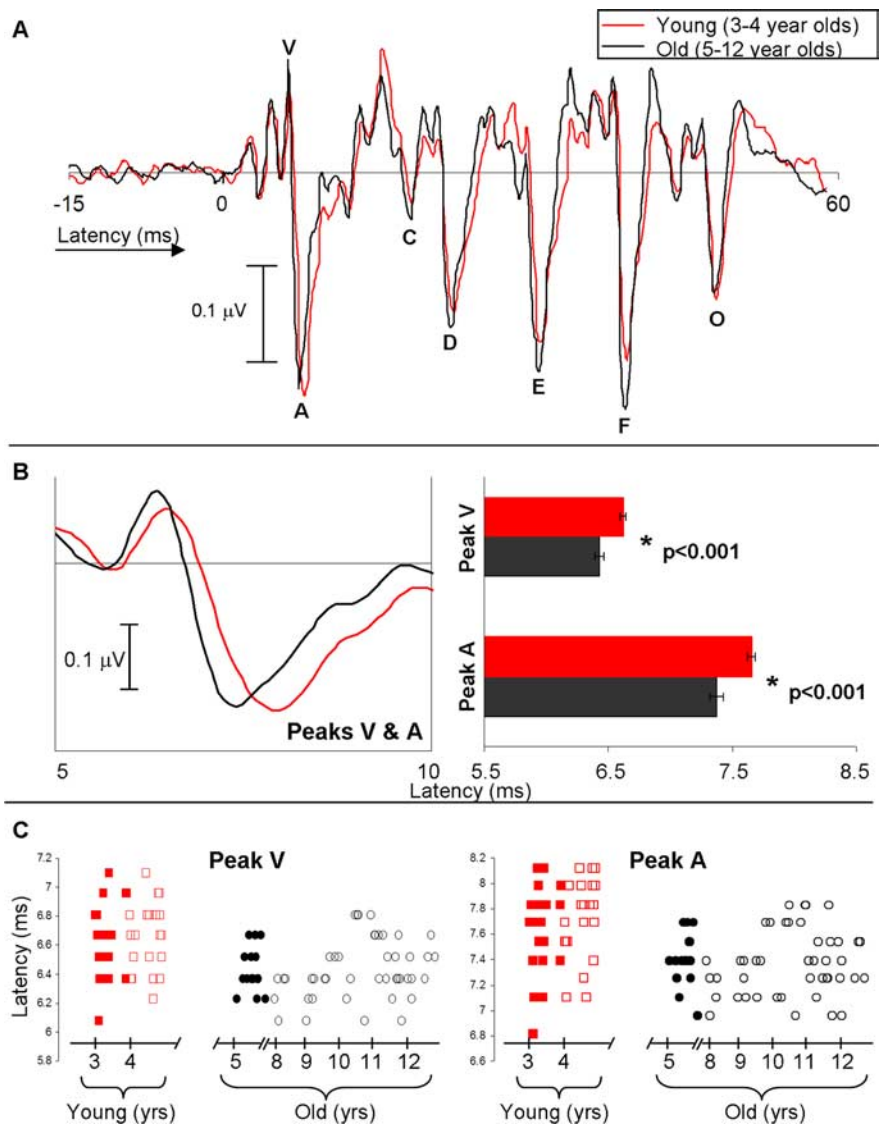Figure 2*A* shows the grand average speech-evoked re-

Figure 2. *A*, Grand average waveform for the young group (3- to 4-year-old subjects) in red and the old group (5- to 12-year-old subjects) in black. *B*, Left, Enlarged VA region of the grand average waveforms. Right, The young group has significantly delayed latencies to peaks V and A compared with the old group ($p < 0.001$, error bars represent 1 SE). *C*, Individual latencies as a function of age/group for all subjects ($n = 104$) for peaks V and A. Red boxes represent the young group (filled symbols are the 3-year-olds), and black circles represent the old group (filled symbols are the 5-year-olds).

Table 1. Least significant difference *p* values for peak V and A latencies between all age groups

	Age (years)	4	5	8–12
V latency	3	0.63	**0.02**	**0.00**
	4		**0.01**	**0.00**
	5			0.92
A latency	3	0.65	**0.04**	**0.00**
	4		**0.02**	**0.00**
	5			0.56

There was no significant difference between 3- and 4-year-olds or between 5- and 8- to 12-year-olds. Bold indicates significance.

sponse for the young and old group. The onset response is a robust positive–negative peak complex occurring at ∼6.5 ms (peaks V and A). The enlarged depiction of this complex and the bar graphs shown in Figure 2*B* display the latency differences between the young and old group for peak V ($F = 20.151$, $p < 0.001$) and peak A ($F = 25.244$, $p < 0.001$). Figure 2*C* displays the

individual peak V and A latencies for each subject. There is no significant correlation between age (in months) and peak V latency (young, $r = 0.193$; old, $r = 0.127$) or peak A latency (young, $r = 0.200$; old, $r = 0.008$). This suggests that the onset latencies do not get gradually earlier as a child approaches the age of 5 years. Additional onset measures that are significant between groups are VA interpeak latency (old group has shorter interpeak latency, $F = 6.770$, $p = 0.011$) and the slope of the VA complex (old group has a steeper slope, $F = 6.594$, $p = 0.012$).

The negative peaks D, E, and F of the waveform in Figure 2*A* (between ~22 and 40 ms) represent peaks that phase lock to the fundamental frequency of the stimulus. There is little literature on the development of the human FFR and none that we are aware of on the development of the speech-evoked FFR. The literature that does exist suggests that infants and adults show similar FFR response properties when elicited by tone bursts (Gardi et al., 1979; Levi et al., 1995). The results of this study indicate that there are significant latency differences between the young and old group with respect to the FFR peaks such that the young group has delayed peak latencies for D, E, and F ($F = 8.235$, $p = 0.005$; $F = 4.270$, $p = 0.041$; $F = 13.520$, $p < 0.001$, respectively) (Fig. 3). Additionally, the young group had significantly reduced peak F amplitude compared with the old group ($F = 5.291$, $p = 0.023$).

Last, the negative peak O at ~48 ms represents the neural response to the offset of the speech sound (Johnson et al., 2007). Again, to our knowledge, there have been no published reports describing the development of the auditory brainstem response to the offset to any stimulus type. In this study, the young group had a significantly delayed peak O latency compared with the old group ($F = 6.250$, $p = 0.014$). It is important to note, however, that although peak O was determined to be a reliable peak based on the criteria set forth in Materials and Methods (>85% reliable detection in the entire subject population), only 82% of the young group displayed a reliable peak O, whereas 95% of the old group did. Thus, we see evidence for delayed development of the morphology of peak O in the young group. To better understand whether latency differences between the young and old group were being inherited from a previous level of processing, latency differences were calculated at each peak (mean of old group minus mean of young group). Latency differences between groups are as follows: peak V, 0.19 ms; peak A, 0.28 ms; peak C, 0.17 ms; peak D, 0.21 ms; peak E, 0.16 ms; peak F, 0.25 ms; and peak O, 0.19 ms. Table 2 shows the mean, SD, and percentage detectability for all transient measures in the young and old groups.

Sustained response measurements
Significant group differences emerged in the overall RMS amplitude, whereby the young group had smaller overall RMS values compared with the old group ($F = 4.060$, $p = 0.047$). To gain a more accurate understanding of these between-group magnitude differences, we conducted an FFT analysis over the periodic portion of the FFR (21.9–40.6 ms) for both groups (Fig. 4). This measurement of the sustained portion of the FFR provides an overall assessment of the magnitude of phase locking to the stimulus fundamental frequency and its harmonics. Average magnitude was computed over 100 Hz bins for each subject to target the response to each harmonic individually. The young group had significantly reduced FFT magnitude for the bins surrounding F_0, H_2, H_3, H_9, and H_{10} ($F = 4.057$, $p = 0.047$; $F = 6.120$, $p = 0.015$; $F = 5.747$, $p = 0.018$; $F = 18.200$, $p < 0.001$; $F = 19.367$, $p < 0.001$, respectively). These findings are illustrated by the bar
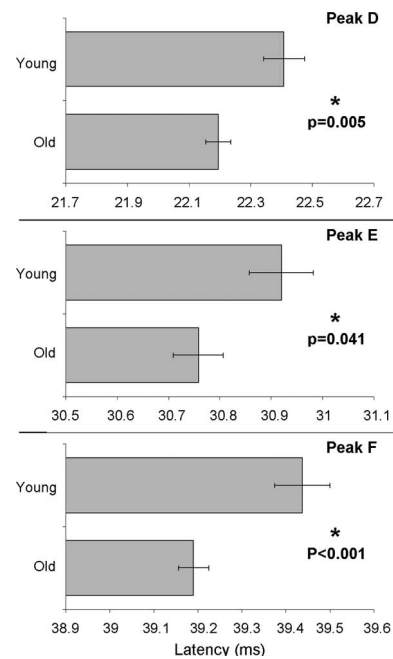

Figure 3. Mean and SE for latencies of FFR peaks D, E, and F. The young group has significantly delayed peak latencies compared with the old group.

graphs of Figure 4. All FFT magnitudes reported were above the noise floor for each group.

Discussion

Ample literature exists to address neural encoding of speech sounds from the eighth nerve (Delgutte, 1980; Sachs and Young, 1980; Miller and Sachs, 1983, 1984), cochlear nucleus (Caspary et al., 1977; Palmer et al., 1986; Keilson et al., 1997; Rhode, 1998; Recio and Rhode, 2000), and brainstem (Galbraith et al., 1995, 1997; Krishnan, 1999, 2002). What motivated the current work was the question of when the encoding of speech develops in the human auditory brainstem. In the present study, evoked potentials were used to analyze the development of the auditory brainstem response to click and speech sounds in children between the ages of 3 and 12 years. The neural response to a click stimulus showed similar response timing across all age groups, in agreement with previously established reports (Salamy, 1984; Gorga et al., 1989; Ponton et al., 1992; Abdala and Folsom, 1995; Hurley et al., 2005). In contrast, peak latency measurements throughout the brainstem response to speech were significantly later for 3- to 4-year-old children compared with 5- to 12-year-olds. Of particular interest is that the onset portion of the speech-evoked response is delayed in the young group, whereas this same portion is equivalent between groups when evoked by a click. This dichotomy suggests that brainstem neurons react differently to encode click versus speech sounds. To our knowledge, this is the first study to show a developmental time course beyond 2 years for encoding stimulus properties in the human brainstem. Furthermore, although studies have addressed how the brainstem FFR responds to speech (Galbraith et al., 1995, 1997; Krishnan, 1999, 2002, 2005; Russo et al., 2004; Johnson et al., 2005), the development of the FFR to speech, which is thought to operate via different mechanisms/pathways than the onset response (Hoormann et al., 1992; Kraus and Nicol, 2005; Song et al., 2006; Akhoun et al., 2008), is unknown. Our data show that latency delays in the young group do not become greater at transient peaks later in the response, consistent with the possibility that latency delays in

the FFR are being inherited from wave V. The extent to which developmental differences in the frequency domain reflect mechanisms that are independent from those observed in the time domain remains to be determined. Thus, although different peaks of the brainstem response can reflect distinct stimulus characteristics such as timing, F_0, and harmonics (Kraus and Nicol, 2005), they may share similar developmental time courses. These data show developmental differences in both onset synchrony and sustained, phase-locked activity. Together, the data support an age-related developmental difference between speech and nonspeech stimuli, in both temporal and frequency domains, and suggest experience-dependent plasticity in the human auditory brainstem.

Click versus speech

Speech is a complex stimulus that, unlike clicks, has environmental relevance and elicits responses that lend themselves to the extraction of information about encoding of syllable onset, offset, and periodicity (pitch). Speech stimuli have a longer rise time and are acoustically more complex compared with clicks. The click stimulus is a short, nonperiodic sound containing a broad range of frequencies, whereas consonant–vowel speech syllables such as /da/ begin with relatively low-amplitude transient onset features followed by a sustained periodic signal, the vowel, which is considerably louder with respect to the consonant. The higher-amplitude, longer-duration vowel may mask the brief consonant, which is critical for eliciting the onset portion of the speech ABR. Backward masking effects have been demonstrated previously in brainstem responses using tone/noise maskers (Marler and Champlin, 2005), and backward masking is known to have a slow developmental time course (Wright and Zecker, 2004; Johnson et al., 2007). Consequently, young children may be more susceptible to neural backward masking effects compared with older children.

Although the acoustic differences discussed above may be partially responsible for the findings in this study, it is important to consider that humans have pervasive exposure to and active engagement with speech, not clicks. Particularly relevant is that brainstem encoding of sound has been shown to be shaped by lifelong linguistic and musical experience (Krishnan et al., 2004, 2005; Musacchia et al., 2007; Wong et al., 2007). That is, brainstem activity evoked by Mandarin tones and music is enhanced in musicians and speakers of tonal languages relative to non-musicians and non-native speakers. Additionally, short-term training has been shown to lead to changes in speech-evoked brainstem activity (Russo et al., 2005; Song et al., 2008). Moreover, recent animal work has shown that experience can lead to large-scale reorganization of the IC tonotopic map (Yu et al., 2007) and that experience-dependent pruning of synaptic inputs is important for the maturation of the functional inhibition in brainstem nuclei (Magnusson et al., 2005). If we assume that humans have little exposure to clicks and that clicks have little relevance, regard-
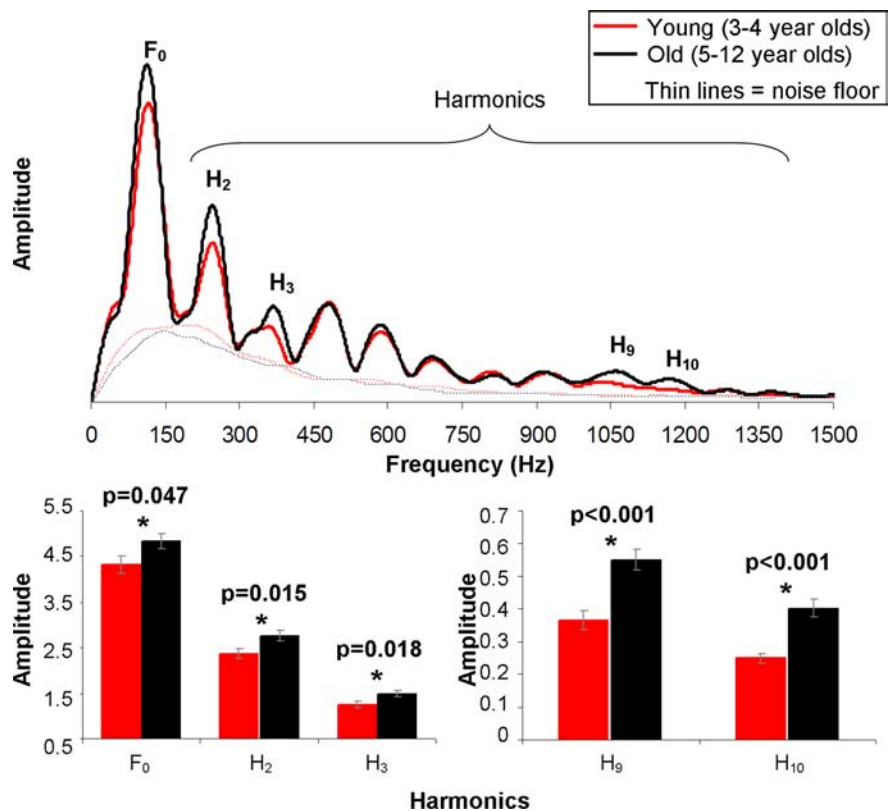

Figure 4. Grand-averaged FFT for the young group (red) and old group (black). Thin lines represent the noise floors for each group. The young group showed significantly diminished magnitudes for the fundamental frequency (F_0) and harmonics 2, 3, 9, and 10. Bar graphs represent mean and SE of significant regions for each group.

less of age, the auditory system would not be expected to change its response to such a stimulus. Conversely, with speech, which is relevant in the real world, experience-dependent pruning is necessary. Because younger children have had less linguistic and phonemic exposure, it is perhaps the case that synaptic pruning has not been fully refined such that young children have delayed/less precise neural response timing when encoding acoustic elements that are relevant to speech. Although it is impossible to answer this question from the data provided here, it is reasonable to speculate that the developmental differences we found may arise not just from acoustic differences but also perhaps from their extensive use and relevance.

Mechanisms for plasticity

This study has identified a developmental time course of speech encoding in the brainstem that suggests neural maturation at the age of 5 years (the age at which most children begin school). At school, the child begins to learn how to read and develops a stronger sense of phonological awareness. Phonological awareness is the ability to identify the different sounds that make words and to associate these sounds with written words to begin reading. It cannot be ruled out that brainstem maturation relevant to encoding speech is a consequence of developing and/or accessing phonological awareness skills. This language-oriented learning may be accompanied by changes in the auditory cortex, similar to those known to be induced by other acoustic experiences (Kilgard and Merzenich, 1998; Kraus and Banai, 2007).

How might cortical or top-down changes influence neural maturation in the brainstem of 5-year-olds? The reverse hierarchy theory suggests that learning modifies the neural circuitry

Table 2. Mean, SD, and percentage detectability for transient measures in the young and old groups

	V		A		VA		C		D		E		F		O	
	Lat*	Amp	Lat*	Amp	Dur*	Slope*	Lat	Amp	Lat*	Amp	Lat*	Amp	Lat*	Amp*	Lat*	Amp
Young (3–4 years)																
Mean	6.62	0.13	7.66	−0.25	1.04	−0.38	18.43	−0.08	22.41	−0.19	30.92	−0.24	39.44	−0.24	48.17	−0.18
SD	0.24	0.05	0.34	0.07	0.21	0.12	0.45	0.05	0.44	0.06	0.40	0.07	0.42	0.09	0.37	0.06
% Detect.	100		100				77		98		95		100		82	
Old (5–12 years)																
Mean	6.43	0.14	7.37	−0.26	0.95	−0.44	18.27	−0.08	22.20	−0.20	30.76	−0.26	39.19	−0.28	47.98	−0.17
SD	0.19	0.05	0.24	0.06	0.08	0.11	0.44	0.05	0.31	0.08	0.37	0.07	0.26	0.09	0.35	0.08
% Detect.	100		100				87		100		95		100		95	

Lat, Latency; Amp, amplitude; Dur, duration; Detect, detectability. Asterisks indicate those measures that displayed significant differences between the young and old groups.

governing performance on a given task starting at the highest level associated with solving the task, gradually refining lower areas when more fine-grained sensory information is required (Ahissar and Hochstein, 2004). Phonological awareness and reading are linked in a bidirectional manner such that phonological awareness facilitates reading ability, and learning to read strengthens phonological awareness skills (Foy and Mann, 2006). One can speculate that the top-down influence of developing phonological awareness skills and reading helps guide plasticity in the auditory brainstem, along with ongoing maturation of the cortex and corticofugal projections. For example, it has been shown that there is a sensitive period for normal cortical maturation between the ages of 3 and 4 years (Sharma et al., 2002). Sharma and colleagues found that congenitally deaf children implanted with a cochlear implant during this sensitive period developed normal cortical responses to sound, whereas those implanted later show substantially altered timing of the cortical response. Our data suggest that, at the level of the brainstem, neural response timing and frequency representation do not develop gradually but are reached somewhat abruptly at the age of 5 years, and this could perhaps be a result of refinement of corticofugal projections with language experience during this cortical sensitive period.

Additional support for experience-dependent plasticity in humans is derived from literature on statistical learning. Although we cannot directly assess the contribution of statistical learning in our data, the literature describes a manner with which the auditory system reacts to frequently occurring sounds. At the level of IC, neural populations rapidly adjust their firing patterns based on the statistical distribution of the sounds encountered, and these adjustments engender improved coding accuracy for sounds occurring most commonly (Dean et al., 2005), even in an on-line manner. Additionally, it has been amply demonstrated that experience shapes the acquisition of many aspects of language-specific sensitivity (Saffran et al., 1996; Jusczyk, 2002). Saffran et al. (1996) demonstrated that infants learn to segment words from fluent speech based on the statistical properties of language input. These studies are consistent with the idea that the human auditory brainstem is susceptible to high-probability, experience-dependent learning/plasticity when encoding sounds composed of acoustic elements relevant to speech until school age.

Last, our findings may provide a biological basis, responsible in part, for the development of phonologic development in children. For example, it is known that children and adults perform differently on perceptual speech identification tasks such as making decisions about voice-onset time, vocalic length and duration, formant transitional periods, and segmental context (Elliott et al., 1986, 1989; Nittrouer, 1996; Mayo and Turk, 2004, 2005). Furthermore, it has been suggested that children's perceptual weighting strategies for speech-relevant acoustic properties change as they gain experience with a native language (Nittrouer and Crowther, 1998), and that the development of metaphonemic awareness may play some role in changes in cue weighting (Mayo et al., 2003). Together, it is reasonable to speculate that immature perception of speech in children may be related to a delayed development of the neural network responsible for accurate encoding of speech acoustics. The results of this study may show neurophysiologic evidence underlying this perceptual phenomenon.

Clinical applications

A growing body of literature has revealed speech-evoked brainstem response differences between normal children and some children with language-based learning problems. The idea that the encoding of linguistic information can serve as a biological marker for auditory function in children with learning and literacy disorders has led to the development of a clinical test (BioMAP) to objectively assess disordered processing of sound in school-age children (8–12 years of age). The present study implies that the age range for such testing can include 5-year-olds. Brainstem responses provide one of few clinical avenues for assessing auditory processing abilities in children as young as 5 years of age, providing a mechanism for early-intervention recommendations and monitoring of educational progress.

References

Abdala C, Folsom RC (1995) The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. J Acoust Soc Am 98:921–930.

Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. Trends Cogn Sci 8:457–464.

Akhoun I, Gallego S, Moulin A, Menard M, Veuille E, Berger-Vachon C, Thai-Van H (2008) The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. Clin Neurophysiol 119:922–933.

Banai K, Nicol T, Zecker SG, Kraus N (2005) Brainstem timing: implications for cortical processing and literacy. J Neurosci 25:9850–9857.

Brainard MS, Knudsen EI (1993) Experience-dependent plasticity in the inferior colliculus: a site for visual calibration of the neural representation of auditory space in the barn owl. J Neurosci 13:4589–4608.

Caspary DM, Rupert AL, Moushegian G (1977) Neuronal coding of vowel sounds in the cochlear nuclei. Exp Neurol 54:414–431.

Cunningham J, Nicol TG, Zecker SG, Kraus N (2001) Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol 112:758–767.

Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. Nat Neurosci 8:1684–1689.

DeBello WM, Feldman DE, Knudsen EI (2001) Adaptive axonal remodeling in the midbrain auditory space map. J Neurosci 21:3161–3174.

Delgutte B (1980) Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. J Acoust Soc Am 68:843–857.

Despland PA, Galambos R (1980) The auditory brainstem response (ABR) is a useful diagnostic tool in the intensive care nursery. Pediatr Res 14:154–158.

Elliott LL, Busse LA, Partridge R, Rupert J, DeGraaff R (1986) Adult and child discrimination of CV syllables differing in voicing onset time. Child Dev 57:628–635.

Elliott LL, Hammer MA, Scholl ME, Wasowicz JM (1989) Age differences in discrimination of simulated single-formant frequency transitions. Percept Psychophys 46:181–186.

Foy JG, Mann V (2006) Changes in letter sound knowledge are associated with development of phonological awareness in pre-school children. J Res Read 29:143–161.

Galbraith GC (1994) Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. Electroencephalogr Clin Neurophysiol 92:321–330.

Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM (1995) Intelligible speech encoded in the human brain stem frequency-following response. NeuroReport 6:2363–2367.

Galbraith GC, Jhaveri SP, Kuo J (1997) Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. Electroencephalogr Clin Neurophysiol 102:46–53.

Gardi J, Salamy A, Mendelson T (1979) Scalp-recorded frequency-following responses in neonates. Audiology 18:494–506.

Gorga MP, Kaminski JR, Beauchaine KL, Jesteadt W, Neely ST (1989) Auditory brainstem responses from children three months to three years of age: normal patterns of response. II. J Speech Hear Res 32:281–288.

Hall JW (1992) Handbook of auditory evoked responses. Boston: Allyn and Bacon.

Hood LJ (1998) Clinical applications of the auditory brainstem response. San Diego: Singular.

Hoormann J, Falkenstein M, Hohnsbein J, Blanke L (1992) The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. Hear Res 59:179–188.

Hurley RM, Hurley A, Berlin CI (2005) Development of low-frequency tone burst versus the click auditory brainstem response. J Am Acad Audiol 16:114–121; quiz, 122.

Irvine DR (1992) Physiology of the auditory brainstem. In: The mammalian auditory pathway: neurophysiology (Popper A, Fay R, eds), pp 153–231. New York: Springer.

Jacobson JT (1985) The auditory brainstem response. San Diego: College-Hill.

Johnson KL, Nicol TG, Kraus N (2005) The brainstem response to speech: a biological marker of auditory processing. Ear Hearing 26:424–434.

Johnson KL, Nicol TG, Zecker SG, Kraus N (2007) Auditory brainstem correlates of perceptual timing deficits. J Cogn Neurosci 19:376–385.

Jusczyk PW (2002) Some critical developments in acquiring native language sound organization during the first year. Ann Otol Rhinol Laryngol Suppl 189:11–15.

Keilson SE, Richards VM, Wyman BT, Young ED (1997) The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: evidence for a periodicity-tagged spectral representation. J Acoust Soc Am 102:1056–1071.

Kilgard MP, Merzenich MM (1998) Plasticity of temporal information processing in the primary auditory cortex. Nat Neurosci 1:727–731.

King C, Warrier CM, Hayes E, Kraus N (2002) Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett 319:111–115.

Klatt DH (1980) Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 67:971–995.

Knudsen EI (1998) Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. Science 279:1531–1533.

Kraus N, Banai K (2007) Auditory-processing malleability: focus on language and music. Curr Dir Psychol Sci 16:105–110.

Kraus N, Nicol TG (2005) Brainstem origins for cortical "what" and "where" pathways in the auditory system. Trends Neurosci 28:176–181.

Krishnan A (1999) Human frequency-following responses to two-tone approximations of steady-state vowels. Audiol Neurootol 4:95–103.

Krishnan A (2002) Human frequency-following responses: representation of steady-state synthetic vowels. Hearing Res 166:192–201.

Krishnan A, Xu Y, Gandour JT, Cariani PA (2004) Human frequency-following response: representation of pitch contours in Chinese tones. Hearing Res 189:1–12.

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

Levi EC, Folsom RC, Dobie RA (1995) Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. Hear Res 89:21–27.

Magnusson AK, Kapfer C, Grothe B, Koch U (2005) Maturation of glycinergic inhibition in the gerbil medial superior olive after hearing onset. J Physiol (Lond) 568:497–512.

Marler JA, Champlin CA (2005) Sensory processing of backward-masking signals in children with language-learning impairment as assessed with the auditory brainstem response. J Speech Lang Hearing Res 48:189–203.

Mayo C, Turk A (2004) Adult-child differences in acoustic cue weighting are influenced by segmental context: children are not always perceptually biased toward transitions. J Acoust Soc Am 115:3184–3194.

Mayo C, Turk A (2005) The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. J Acoust Soc Am 118:1730–1741.

Mayo C, Scobbie JM, Hewlett N, Waters D (2003) The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. J Speech Lang Hear Res 46:1184–1196.

Miller MI, Sachs MB (1983) Representation of stop consonants in the discharge patterns of auditory-nerve fibers. J Acoust Soc Am 74:502–517.

Miller MI, Sachs MB (1984) Representation of voice pitch in discharge patterns of auditory-nerve fibers. Hearing Res 14:257–279.

Moore JK, Linthicum Jr FH (2007) The human auditory system: a timeline of development. Int J Audiol 46:460–478.

Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci USA 104:15894–15898.

Nittrouer S (1996) Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. J Speech Hear Res 39:278–297.

Nittrouer S, Crowther CS (1998) Examining the role of auditory sensitivity in the developmental weighting shift. J Speech Lang Hear Res 41:809–818.

Palmer AR, Winter IM, Darwin CJ (1986) The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons. J Acoust Soc Am 79:100–113.

Ponton CW, Eggermont JJ, Coupland SG, Winkelaar R (1992) Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: evidence from derived auditory brain-stem responses (ABRs). J Acoust Soc Am 91:1576–1586.

Recio A, Rhode WS (2000) Representation of vowel stimuli in the ventral cochlear nucleus of the chinchilla. Hear Res 146:167–184.

Rhode WS (1998) Neural encoding of single-formant stimuli in the ventral cochlear nucleus of the chinchilla. Hear Res 117:39–56.

Russo N, Nicol TG, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. Clin Neurophysiol 115:2021–2030.

Russo N, Nicol TG, Zecker SG, Hayes EA, Kraus N (2005) Auditory training improves neural timing in the human brainstem. Behav Brain Res 156:95–103.

Sachs MB, Young ED (1980) Effects of nonlinearities on speech encoding in the auditory nerve. J Acoust Soc Am 68:858–875.

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274:1926–1928.

Salamy A (1984) Maturation of the auditory brainstem response from birth through early childhood. J Clin Neurophysiol 1:293–329.

Sharma A, Dorman MF, Spahr AJ (2002) A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. Ear Hear 23:532–539.

Song JH, Banai K, Russo NM, Kraus N (2006) On the relationship between speech- and nonspeech-evoked auditory brainstem responses. Audiol Neurootol 11:233–241.

Song JH, Skoe E, Wong PC, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci, in press.

Wible B, Nicol T, Kraus N (2004) Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. Biol Psychol 67:299–317.

Wible B, Nicol T, Kraus N (2005) Correlation between brainstem and cortical auditory processes in normal and language-impaired children. Brain 128:417–423.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 10:420–422.

Wright BA, Zecker SG (2004) Learning problems, delayed development, and puberty. Proc Natl Acad Sci USA 101:9942–9946.

Xu Y, Krishnan A, Gandour JT (2006) Specificity of experience-dependent pitch representation in the brainstem. NeuroReport 17:1601–1605.

Yu X, Sanes DH, Aristizabal O, Wadghiri YZ, Turnbull DH (2007) Large-scale reorganization of the tonotopic map in mouse auditory midbrain revealed by MRI. Proc Natl Acad Sci USA 104:12193–12198.

Zhang Y, Suga N (2005) Corticofugal feedback for collicular plasticity evoked by electric stimulation of the inferior colliculus. J Neurophysiol 94:2676–2682.

Zheng W, Knudsen EI (1999) Functional selection of adaptive auditory space map by GABAA-mediated inhibition. Science 284:962–965.