Auditory Brainstem Correlates of Perceptual Timing Deficits

Krista L. Johnson, Trent G. Nicol, Steven G. Zecker, and Nina Kraus

Abstract

■ Children with language-based learning problems often exhibit pronounced speech perception difficulties. Specifically, these children have increased difficulty separating brief sounds occurring in rapid succession (temporal resolution). The purpose of this study was to better understand the consequences of auditory temporal resolution deficits from the perspective of the neural encoding of speech. The findings provide evidence that sensory processes relevant to cognition take place at much earlier levels than traditionally believed. Thresholds from a psychophysical backward masking task were used to divide children into groups with good and poor temporal resolution. Speech-evoked brainstem responses were analyzed across groups to measure the neural integrity of stimulus-time mechanisms. Results suggest that children with poor temporal resolution do not have an overall neural processing deficit, but rather a deficit specific to the encoding of certain acoustic cues in speech. Speech understanding relies on the ability to attach meaning to rapidly fluctuating changes of both the temporal and spectral information found in consonants and vowels. For this to happen properly, the auditory system must first accurately encode these time-varying acoustic cues. Speech perception difficulties that often co-occur in children with poor temporal resolution may originate as a neural encoding deficit in structures as early as the auditory brainstem. Thus, speech-evoked brainstem responses are a biological marker for auditory temporal processing ability. ■

INTRODUCTION

It is estimated that between 5% and 10% of otherwise unimpaired children are afflicted with some type of a language-based learning problem (LP) such as dyslexia, central auditory processing deficit, language-learning impairment, and/or attention deficit hyperactivity disorder (Shaywitz, 1998; Torgeson, 1991). Parts of this heterogeneous LP population have a pronounced difficulty detecting rapid changes in sound over time, or poor temporal resolution (Wright et al., 1997; Elliott & Hammer, 1993; Tallal, Miller, & Fitch, 1993; Elliott, Hammer, & Scholl, 1989). Likewise, these children with LPs also demonstrate deficient neural encoding of speech in the brainstem (Banai, Nicol, Zecker, & Kraus, 2005; Johnson, Nicol, & Kraus, 2005; Wible, Nicol, & Kraus, 2004; King, Warrier, Hayes, & Kraus, 2002; Cunningham, Nicol, Zecker, & Kraus, 2001). Although many researchers have speculated that some LP children have an underlying auditory perceptual deficit that results in a fundamental inability to hear acoustic distinctions in speech (Kraus et al., 1996; Elliott et al., 1989; Tallal, Stark, & Mellits, 1985; Tallal, 1980), few attempts have been made to link these perceptual deficits to the underlying neural processes involved with low-level auditory pathway encoding. Here we investigate the relation between perceptual and neural timing deficits associated with speech encoding.

This idea dovetails with recent models of how the speech-evoked brainstem response encodes the acoustic source and filter cues inherent in speech (Johnson et al., 2005; Kraus & Nicol, 2005; Fant, 1960). Generally, source cues convey information important for talker identity and prosodic content, whereas filter cues convey information about the specific consonants and vowels spoken, important for understanding the linguistic ''message.'' According to this framework, discrete transient peaks within the brainstem response are ''responsible'' for encoding acoustic information simultaneously conveyed in speech. The encoding of source information (fundamental frequency, F_0) can be distinguished from the neural encoding of filter cues (mainly acoustic onset, harmonics of F_0, and acoustic offset). Source and filter characteristics have, in turn, been linked to cortical ''what'' and ''where'' pathways (Kraus & Nicol, 2005). Because it is thought that a basic deficit in the processing and integration of rapidly successive and transient signals might be to blame for the auditory perceptual deficit experienced by many LP children (Merzenich et al., 1996; Tallal et al., 1996), this conceptual framework of neural encoding was used to evaluate the rapid processing hypothesis, which states that deficits in processing transient rapid acoustic signals weaken the ability to distinguish phonemes (Temple et al., 2000).

Northwestern University

Previous work has touched on the association between poor temporal resolution and neural encoding deficits. A recent study found that LP children have brainstem neural delays in conditions evoked by nonspeech backward masking stimuli (Marler & Champlin, 2005). This is important because it demonstrates that LP children are more susceptible to backward masking at the neural level as well as the behavioral level, but does not establish the link between poor temporal resolution and speech encoding. Other studies have begun to describe speech-evoked brainstem encoding deficits in children with language-based LPs (Banai et al., 2005; Wible et al., 2004; King et al., 2002; Cunningham et al., 2001). A consistent finding is that about one third of LP children exhibit delayed peak latency and/or shallower slope measures of the onset response, while also displaying diminished first-formant (F_1) spectral magnitude compared with normal learning (NL) children. These deficits exist with little evidence of F_0 encoding problems. The purpose of the current study is to investigate the relationship between perceptual temporal resolution and the neural encoding of speech sounds. Temporal resolution is measured by a psychophysical backward masking task based on the established and pronounced difficulties some LP children have been shown to exhibit (Wright et al., 1997). It should be noted that backward masking stimuli used in this study and the acoustic characteristics of stop consonants in consonant–vowel syllables share parallel characteristics. Both contain a rapid and brief onset of relatively small amplitude followed by a much longer and louder segment. We evaluated whether LP children with abnormal perceptual temporal resolution differ from NL children with good temporal resolution in their brainstem representation of acoustic elements of speech. Results are interpreted within the context of the source-filter model of speech acoustics and the neural representation of these signal attributes (Johnson et al., 2005; Kraus & Nicol, 2005; Fant, 1960) with the expectation that this will inform our understanding of the unique neural underpinnings associated with perceptual timing deficits, as they relate to speech. We hypothesized that poor performance on a temporal resolution task, backward masking, is directly related to the deficient neural encoding of speech sounds at the level of the brainstem, and that neural encoding deficits would occur for the filter rather than source characteristics of speech. This would explain, in part, why temporal resolution deficits so often accompany language-based LPs, as cognitive tasks such as literacy appear to be affected by sensory function, and vice versa.

METHODS

Subjects

Seventy children between the ages of 8 and 12 years participated in this study. All children had normal pure tone thresholds as assessed by screening at 20 dB HL for octaves from 500 to 4000 Hz, and normal wave V click auditory brainstem response (ABR) latencies. Forty-six children (13 girls, 33 boys) were classified as having LPs based on diagnosis by an outside professional and a study-specific psychoeducational test battery.[1] Twenty-four children (10 girls, 14 boys) were classified as normal learning (NL) based on subject history and the psychoeducational test battery. By definition, a person with a learning disability has normal or above normal IQ, with performance in one or more areas of achievement (i.e., math, reading, spelling) being significantly lower than expected based on IQ. Therefore, IQ, as measured by the Brief Cognitive Scale or Test of Nonverbal Intelligence-3 (TONI-3; Sherbenou & Johnson, 1997) scores, was greater than 85 for each individual in both NL and LP groups. Finally, none of the subjects had low birth weight, nor was there a between-group difference with regard to premature births (linear-by-linear association, $p = .235$) as reported by their parents/guardians in a subject history form. Written and oral informed consent was given by each child and his or her parent or guardian. The children were paid for their participation. Institutional review board approval for this study was obtained from Northwestern University.

Psychophysics

Stimuli

The digitally generated stimuli consisted of a signal and a masker. The signal was a 1000-Hz, 20-msec pure tone immediately followed by a 300-msec noise band-passed from 600 to 1400 Hz. Both the tone and noise stimuli included a 10-msec cosine squared gating envelope. The signal was presented 20 msec before the onset of the masker and ended with the beginning of the masker. Stimuli were presented binaurally with Sennheiser HD 540 headphones. Stimuli were identical to those used by Wright et al. (1997).

Procedure

A standard, adaptive, two-interval forced-choice (2AFC) maximum likelihood technique was used to estimate the tone level required for 94% correct detection (Green, 1990). Each trial consisted of two intervals: one containing the tone–noise pair; one containing just noise. Observation intervals were separated by 800 msec. The initial intensity of the signal was 99 dB SPL and the masker remained at a constant level of 40 dB SPL. Subjects were instructed to press the key that corresponded to the observation interval in which they heard the signal. Visual feedback was given from a computer screen that also marked the observation intervals. For each trial, the tone level was determined by a psychometric function that was estimated to have produced the responses up to that point in the test. A total of

three threshold measurements were obtained for each subject, and the final threshold was calculated as the average of the second and third threshold estimates to rule out effects of procedural learning.

Electrophysiology

Stimuli

A Klatt cascade/parallel formant synthesizer (Klatt, 1980) was used to synthesize the speech stimulus /da/ at a sampling rate of 10 kHz. The syllable was 40 msec in duration, and consisted of five formants with an onset burst frication during the first 10 msec at F_3, F_4, and F_5. Stimuli were presented with randomly alternating polarity by a PC-based stimulus delivery system (NeuroScan Gentask, Compumedics USA, El Paso, TX) that output the signals through a 16-bit converter. At stimulus onset, a PC-based evoked potentials averaging system (NeuroScan Acquire, Compumedics USA, El Paso, TX) was triggered. 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 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, with the implication being that behavioral temporal processing is mediated by how the brain represents the acoustic sound structure at an automatic and preconscious level.

Physiological Recording

Recording brainstem response to sound has long been established as a valid and reliable means to assess the integrity of the neural transmission of acoustic stimuli. Transient acoustic events induce a pattern of voltage fluctuations in the brainstem resulting in a familiar waveform, yielding information about brainstem nuclei along the ascending central auditory pathway (for a review, see Hood, 1998; Jacobson, 1985). An accurate manifestation of stimulus timing in the auditory brainstem is a hallmark of normal perception (Sininger & Starr, 2001). Brainstem responses were collected according to widely used procedures as described in detail by Hood (1998) and Jacobson (1985). All recordings were made with silver–silver chloride electrodes (impedance <5 kΩ). Because the deep generator sources of brainstem responses make it such that electrode position will minimally affect the response measured, responses were differentially recorded from Cz-to-ipsilateral earlobe, with forehead as ground. Three blocks of 2000 responses were collected at a rate of 10.99/sec. A 70-msec recording window (including a 10-msec prestimulus period) was used. Responses were sampled at 20,000 Hz and bandpass filtered online from 100 to 2000 Hz. Sweeps with activity exceeding ±35 μV were rejected from the average. A total of 6000 artifact-free responses were used to create an averaged response for each subject.

Brainstem Response Analysis

In order to isolate the neural response from that of the cochlear microphonic and eliminate stimulus artifact, responses to the alternating polarities were added together (Gorga, Abbas, & Worthington, 1985). The brainstem response to the speech sound /da/ can be described in terms of both discrete peaks and a sustained response. The discrete peaks are transient responses that exist for only tenths of a millisecond. Transient peaks with published norms consist of the onset response (waves V and A) and additional peaks following the onset (waves C and F) (Russo, Nicol, Musacchia, & Kraus, 2004; King et al., 2002). In this study, additional transient peaks D, E, and O were evaluated. Transient responses provide information relevant to both neural timing and magnitude. Peak latency and amplitude were determined for all waves. The VA onset complex was further analyzed by computing slope.

The sustained, frequency-following response (FFR) to /da/ lasts for about 35 msec and encodes the ongoing harmonic information within the speech syllable. This region is represented in the response from the conclusion of the onset response to approximately 48 msec and was analyzed using three measures. First, the root-mean-square (RMS) amplitude was calculated over the time range of 12.5–47.5 msec. This was used to quantify the overall magnitude of the sustained activity, providing a measure of an individual's neural population response. Second, stimulus-to-response (SR) cross-correlations and lags were calculated over the 10- to 40-msec period of the stimulus. SR correlations examine how well the response maintains the morphological features of the stimulus. Cross-correlation was used to shift the response in time until a maximum correlation with the stimulus was reached. The SR lag represents the amount of delay between the stimulus and response at maximum correlation. This measure provides information regarding phase locking of the response to the crucial frequencies present in the speech stimulus. Lastly, fast Fourier transform (FFT) analysis of the response was performed over the period of 12.5–47.5 msec to evaluate the spectral composition of the response. The magnitudes of frequency representation over the stimulus F_0 (103–171 Hz), F_1 (220–720 Hz), and a higher frequency portion of F_1 (400–650 Hz) were measured by taking the average of the amplitudes over the specified frequency ranges.

RESULTS

There is a known developmental time course for backward masking. Adultlike performance is not reached until

approximately 10–11 years of age (Hartley, Wright, Hogan, & Moore, 2000). Consequently, these data show a significant correlation between age and backward masking score ($r = -.551$, $p = .027$). To avoid developmental variability, subjects were divided into two age groups: young (8–10 years old; NL, $n = 9$; LP, $n = 25$) and old (11–12 years old; NL, $n = 15$; LP, $n = 21$) as per Hartley et al. (2000). Normative data from our laboratory have established the mean backward masking threshold for the younger children to be 64.72 dB SPL (*SD* = 14.75) compared to 52.88 dB SPL (*SD* = 16.08) for the older children. Backward masking thresholds were significantly different between these age groups (analysis of variance [ANOVA]: $F = 14.189$, $p = .001$). Normal ranges for each age group were calculated by the mean ± 1.5 *SD* (42.59–86.85 and 28.76–77.01 dB SPL for the young and older subjects, respectively). All NL subjects in this study ($n = 24$) had backward masking thresholds within the normal range.

Backward masking thresholds revealed a number of LP subjects who had thresholds greater than their age-appropriate normal values (44% of the young LP subjects and 24% of the old LP subjects), as shown in Figure 1. Children with poor backward masking scores ($n = 16$) were combined into one group and defined as the LP-out group. The remainder of the LP children ($n = 30$) were classified as the LP-in group. A two-way ANOVA with backward masking threshold as the dependent variable showed no interaction between age and group, demonstrating that developmental shifts in backward masking thresholds across groups are comparable ($F = 1.533$, $p = .220$). Furthermore, a chi-square indicated there is no significant difference in the proportion of young versus old children among the three groups ($\chi^2 = 3.829$, $p = .147$).

In order to understand the early neurophysiological timing mechanisms associated with backward masking, the brainstem responses were analyzed with respect to these three groups. In no case are the following neurophysiological findings attributed to maturational effects of the brainstem response. The auditory brainstem response to clicks and tones is mature by age 5 (Inagaki et al., 1987; Salamy, 1984), and preliminary work suggests the speech-evoked brainstem response is mature by age 3 (Johnson, Nicol, & Kraus, 2006). There were no significant differences between any of the neurophysiological measures and age among any of the three groups (ANOVA: $p > .05$). Moreover, IQ was not a contributing factor, as the significant between-group differences remain stable when the effect of IQ is removed by treating it as a covariate (analysis of covariance: $p < .05$ for all neurophysiological measures).

Transient Peak Latency and Amplitude Measures

Table 1 gives mean and standard deviation values for all transient peak measures in each group. There were significant between-group differences for wave A, C, and O latencies (ANOVA: $F = 3.412$, $p = .039$, $F = 7.019$, $p = .002$, $F = 14.057$, $p < .001$, respectively). Least significant difference (LSD) post hoc analyses showed no significant differences between NL and LP-in ($p > .05$) for any of the measures. Post hoc tests revealed that the LP-out group had significantly delayed latencies for waves A, C, and O compared to NL ($p = .019$, $p < .001$, $p < .001$, respectively) and LP-in ($p = .024$, $p = .011$, $p < .001$, respectively). Figure 2A illustrates these findings. There were no significant amplitude differences for any of the peaks ($p > .05$). The latencies of waves V, D, E, and F did not differ among groups ($p > .05$). Interestingly, it appears that the latency discrepancies between waves A, C, and O in normal and LP-out children become progressively larger at subsequent peaks. The mean delay between NL and LP-out children was 0.24 msec at wave A, 0.38 msec at wave C, and 0.62 msec at wave O, with

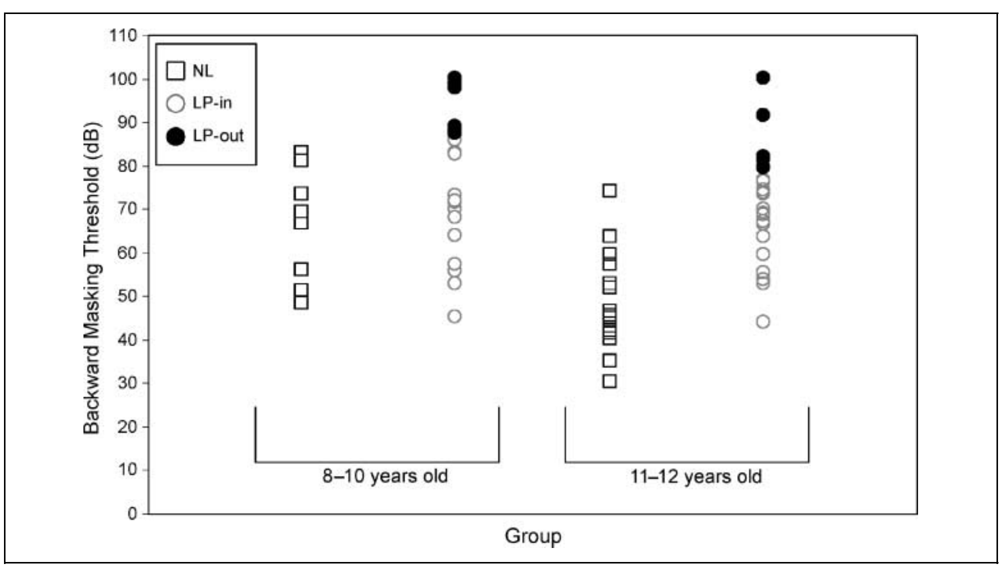

Figure 1. Backward masking thresholds for NL children (squares) and children with LP (circles). Because backward masking thresholds are age dependent, the children were divided into two age groups (8–10 and 11–12 years old). The normal range for each age group was determined by the mean ± 1.5 *SD* of normal ranges. LP children who fell outside this range (filled circles) were analyzed as a separate group.

Table 1. Peak Latency Means and Standard Deviations for NL and the Two LP Groups

	V	*A*	*C*	*D*	*E*	*F*	*O*	*VA Slope*
NL								
M	7.78	8.64	18.80	23.26	32.06	40.60	48.86	1.16
SD	0.25	0.26	0.35	0.46	0.36	0.37	0.43	0.33
LP-in								
M	7.76	8.66	18.95	23.27	32.14	40.55	49.08	1.12
SD	0.23	0.28	0.37	0.27	0.37	0.31	0.43	0.39
LP-out								
M	7.87	8.86	19.27	23.46	32.30	40.67	49.65	0.90
SD	0.28	0.33	0.46	0.35	0.41	0.46	0.57	0.23

correspondingly increasing effect sizes, 0.74, 1.15, and 1.57, respectively. Furthermore, a discriminant analysis suggested that among peaks A, C, and O, peak O latency provides the strongest prediction of poor temporal resolution (82.5% of original group cases were classified correctly compared to chance level of 50%). However, the significant latency delay of peak O is attributable to more than just the cumulative effects of a late peak A and peak C. That is, there remains a significant between-group difference for peak O when the effects of the latencies of peaks A, C, D, E, and F are removed by treating them as covariates ($F = 9.608$, $p < .01$).

The slope of the VA onset response also demonstrated significant group differences (ANOVA: $F = 3.240$, $p = .045$), with LSD post hoc analysis showing a broader slope for the LP-out group compared to NL ($p = .018$) and LP-in ($p = .037$) groups.

Sustained Response Measurements

There were no significant group differences for RMS amplitude or SR correlation and lag (ANOVA: $p > .05$). There were no differences in FFT magnitude across the F_0 frequency range ($p > .05$). When the entire frequency range encompassed by F_1 was considered (220–720 Hz) no group differences emerged. However, there were significant group differences for the 400- to 650-Hz range (ANOVA: $F = 3.424$, $p = .038$). LSD post hoc analysis showed no difference between the NL and LP-in groups ($p > .05$). The LP-out group showed decreased 400- to 650-Hz magnitude compared to the NL group ($p = .015$) but not the LP-in group ($p > .05$). Figure 2B illustrates this difference.

DISCUSSION

This study established a relationship between backward masking thresholds and the neurophysiological brainstem response to a speech syllable. To our knowledge, this is the first study to describe how perceptual temporal resolution deficits affect the neural encoding of speech at the level of the brainstem. These findings provide evidence that sensory processes relevant to cognition take place at much earlier levels than traditionally believed. Learning-impaired individuals with poor backward masking thresholds showed neurophysiological timing deficits in both transient and sustained brainstem responses to a speech sound. A significant number of the LP children (LP-out group) had elevated backward masking thresholds compared to NL children. For all of the brainstem response measures analyzed, the NL and LP-in groups did not differ from one another. By comparison, the LP-out group displayed a number of significant differences, particularly in transient response measures. Analyses revealed a unique set of similarities among, and differences between, the LP-out group and the other groups, most notably deviant onset and offset responses with normal transient peaks within the FFR. The LP-out group also showed diminished spectral energy in the high-frequency range contained in F_1. Although it would be interesting to look at frequency responses above F_1, the phase-locking capabilities of the brainstem prohibit this (Joris, Schreiner, & Rees, 2004; Frisina, 2001; Wang & Sachs, 1994; Frisina, Smith, & Chamberlain, 1990; Blackburn & Sachs, 1989; Langner & Schreiner, 1988).

Relationships between Backward Masking and Speech Perception

Accurate temporal processing is essential in decoding speech. The backward masking findings reported here reinforce previous research, linking children with LPs with poor backward masking thresholds (Marler, Champlin, & Gillam, 2002; Wright et al., 1997).

Backward masking provides an index of an individual's perception of rapidly occurring acoustic stimuli. Although not speech, the backward masking task is temporally similar to many speech sounds in the English language. Consider, for example, any stop consonant–vowel combination, such as /da/. The onset /d/ creates a spectral burst that is unique only to that sound and acoustically cues the listener to important information about place of articulation (Stevens & Blumstein, 1978). Following the onset burst is a more periodic and much louder voiced vowel /a/. The transition between the onset and the periodic sound is an essential cue for the listener to determine information about the preceding consonant. In the backward masking stimulus, a tone is followed immediately by a masking noise; the tone is analogous to the onset burst, whereas the masker is analogous to the transition to the vowel. If an individual has poor temporal resolution when performing a backward masking task, this may translate into a poor ability to correctly separate the onset burst from the periodic

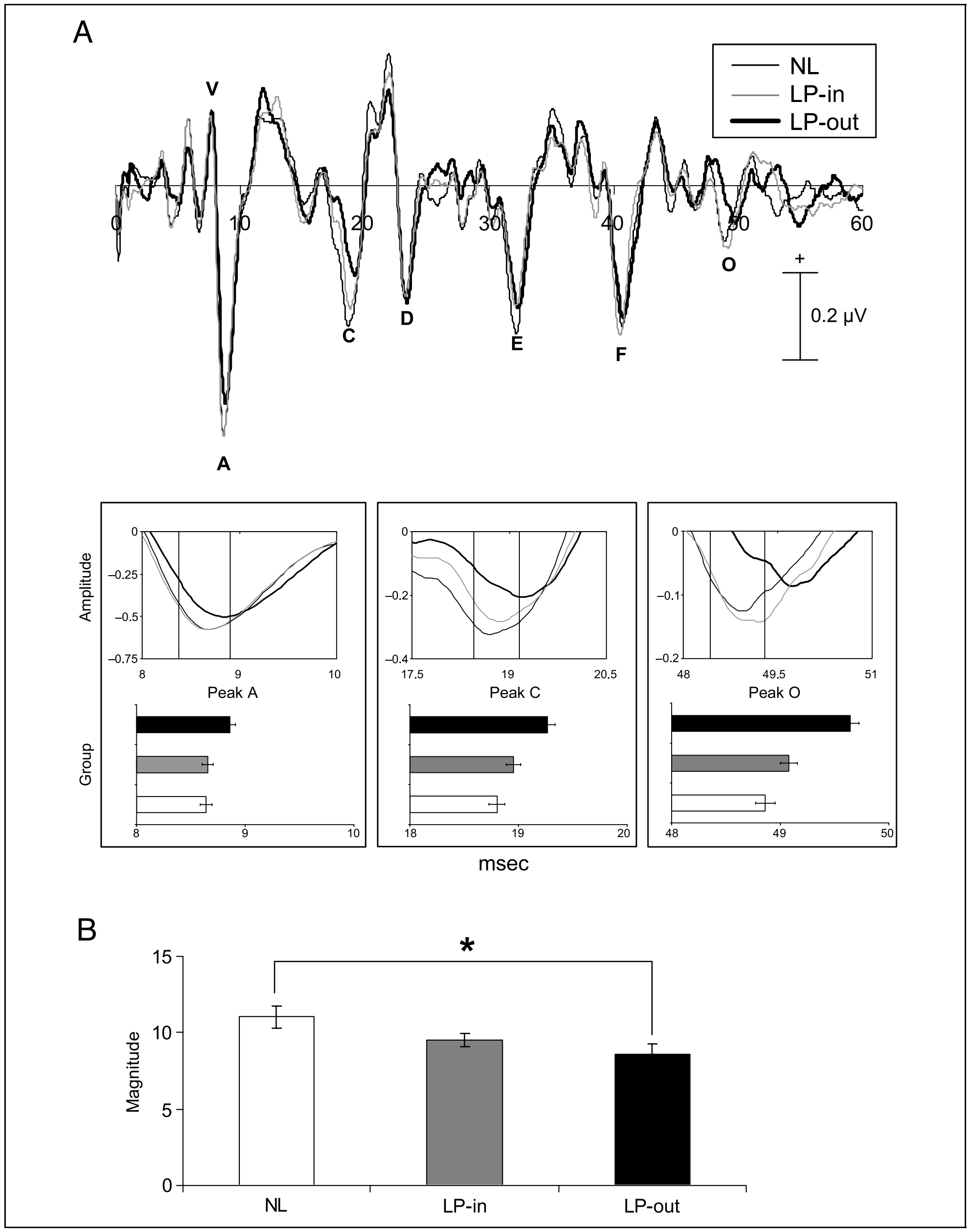

Figure 2. (A) Grand-averaged waveforms for all groups. The LP-out group had significantly delayed latencies for waves A, C, and O compared to NL and LP-in children. There were no between-group latency differences for peaks D, E, and F. (B) Average magnitudes of the response spectra for a high-frequency portion of F_1 (400–650 Hz) across all groups. LP-out children have a decreased magnitude as compared to NL children.

component in speech sounds and misinterpreting the appropriate transitions in ongoing speech, thus leading to poor consonant perception.

Characteristics of the Auditory Brainstem Response: Encoding of Onset, Offset, and Harmonic Components of the Stimulus

The timing of the brain's response to acoustical stimuli can be precisely measured electrophysiologically, and subcortical responses to various acoustical stimuli have long been regarded as reliable means for evaluating the integrity of neuronal timing mechanisms; deviations within fractions of milliseconds are clinically significant in the diagnosis of hearing loss and brainstem function (Hood, 1998). In the properly functioning auditory system, the presence of sound stimulates neuronal activity that is synchronized to the onset of the sound. The onset response to a consonant–vowel speech syllable is reflected by the transient VA complex of the brainstem response. Following this onset response is the sustained FFR that represents a neural population that phase locks to the periodic aspects of the syllable (Worden & Marsh, 1968). This sustained response contains transient peaks synchronized to the fundamental and harmonic frequency components of the stimulus (Wilson & Dobie, 1987; Dobie & Wilson, 1984; Sohmer, Pratt, & Kinarti, 1977). A less commonly discussed attribute of neural response to acoustical stimuli is the offset response. Offset responses in the auditory brainstem response were first

described by Brinkmann and Scherg (1979). Compared to the wave V onset, they described a reversed polarity response to the offset of a Gaussian noise burst, thought to originate from a synchronized decrease in auditory nerve firing rate. More recently described is the notion of duration-tuned neurons that may underlie the far-field offset response. Duration-tuned neurons, which usually respond to offsets (Faure, Fremouw, Casseday, & Covey, 2003), are first present in the inferior colliculus (IC) and are thought to respond through the interaction of temporally offset excitatory and inhibitory inputs (Casseday, Ehrlich, & Covey, 2000).

Similar to the results of Banai et al. (2005), Wible et al. (2004), King et al. (2002), and Cunningham et al. (2001), this study demonstrated that auditory brainstem responses to speech onset differ between normal children and some children with learning disabilities. This study further demonstrated reduced phase locking of the FFR in some children with language-based LPs, another previously reported result (Cunningham et al., 2001). Also in agreement with King et al. is the finding that some LP children have delayed wave C latencies. Two new findings have emerged in this subpopulation of LP children with poor temporal resolution: delayed offset responses and stable temporal coding of transient responses within the FFR.

A possible explanation for the latency delay in waves A and C observed in the LP-out group is an overall neural timing delay with latency shifts that propagate through the brainstem. However, evidence suggests that this is not the case. Waves D, E, and F that follow A and C showed no variation between groups, yet the later wave O was again found to display a latency delay.

Taken as a whole, the data in this study support the model for how speech sound structure is represented by brainstem neurons presented by Johnson et al. (2005) and Kraus and Nicol (2005). It appears that waves A, C, and O are generated by neural mechanisms reflecting the transient characteristics associated with filter aspects of speech, whereas waves D, E, and F are generated by neural mechanisms that encode the source information such as F_0. The peak-to-peak latencies between waves D, E, and F faithfully represent frequencies (calculated as 1/period) within the range of the fundamental (103–125 Hz), and are thus thought to represent the FFR. Conversely, the period between waves C and D does not follow this pattern, representing a higher frequency than the fundamental, implicating a transient/filter encoding role for wave C.

Temporal, frequency, and duration information about a stimulus are encoded early in the auditory pathway. The auditory brainstem encodes the onset of the acoustic signal as well as phase locks to the periodic frequency components of the stimulus. Neurons responsible for encoding these properties are thought to reside in the IC (Møller & Jannetta, 1983; Smith, Marsh, & Brown, 1975). Also present in the IC are neurons tuned to signal duration, with different cells having different best durations (Ehrlich, Casseday, & Covey, 1997; Casseday, Ehrlich, & Covey, 1994). A signature property of duration-specific neurons is that they respond to the offset of a stimulus. It is thus feasible that wave C is responding to the offset of the initial burst, and wave O is responding to the offset of the entire stimulus.

Encoding Frequency Components of the Stimulus

Just as measuring transient peak amplitudes and latencies gives information about the synchronous firing of neuronal ensembles, measurements of the sustained portion of the FFR provide an overall assessment of the magnitude of phase locking to the stimulus fundamental frequency and its harmonics. The FFT analysis demonstrated no group differences in the fundamental frequency range. This is supported by the finding that the stimulus-to-response correlations over this range showed no group differences, indicating that the major portions of the stimulus reflected in the response (i.e., the peaks coding the fundamental frequency) were intact. Although there were also no differences in FFT magnitude for the overall first formant frequency range, differences did arise in FFT magnitude over the higher frequency portion of the F_1 range (400–650 Hz). One possible explanation for this is that temporal processing deficits adversely affect frequency representation. The fact that the RMS amplitude over the FFR range did not differ between groups suggests that there is no difference in the overall magnitude of the neuronal population responding to the speech sound between groups. Rather, it appears that the temporal deficits seen in the LP-out group indicate a limited phase-locking capacity specific to higher frequencies.

Backward Masking and its Biological Origins

What role might the subcortical auditory pathway play in backward masking? As this study indicates, LP children with very poor backward masking thresholds have brainstem timing deficits. The auditory brainstem response pattern associated with excessive auditory backward masking is consistent with known physiological aspects of backward masking. Single-cell recordings in the cat indicate that backward masking occurs in the IC but does not appear below the midbrain (Watanabe, 1978). Furthermore, temporal masking experiments modeling backward masking by studying duration-tuned neurons of the IC have shown that nonexcitatory suppression tones can evoke an inhibitory response that suppresses excitatory neural activity to the previous best duration tone (Faure et al., 2003). Because duration-tuned neurons respond to the offset of a stimulus, the abnormal offset responses observed here may be associated with abnormal neural inhibitory properties.

Why might children who perform poorly on a backward masking task display abnormal aggregate neural responses to speech? Perhaps the high-intensity sound following the onset burst in the stimulus creates a backward masking effect on that burst, thus disturbing the normal offset response. Abnormal neurons that encode temporal information (whether it be onset or offset) in the IC or nearby structures could be responsible for the transmission of improper timing cues, thus impairing both physiological and behavioral temporal representation of acoustical cues. Interestingly, these data show that waves C and, in particular, O, which are thought to encode acoustic offsets, were the most significantly different measures between groups. In fact, the latency delay displayed by the LP-out children for peak O cannot simply be attributed to the cumulative effects of delayed peaks earlier in the response. This is an intriguing finding when thought of in relation to the backward masking task. Good performance on the backward masking task requires, in part, the ability to perceive the offset of one sound to be different from the onset of another sound. If the neural mechanism responsible for encoding acoustic offset information is impaired, poor temporal processing is likely to ensue. Based on these data, it is reasonable to surmise that a prominent deficiency in LP children with poor behavioral temporal processing is improper neural encoding of acoustic offsets.

Cortical Consequences

Strong associations between brainstem timing mechanisms and cortical function have been established (Abrams, Nicol, Zecker, & Kraus, 2006; Banai et al., 2005; Wible, Nicol, & Kraus, 2005). Accurate brainstem encoding of the onset of speech is essential for robust processing at the cortical level (Wible et al., 2005), normal patterns of hemispheric dominance (Abrams et al., 2006), as well as cortical sensitivity to acoustic change (Banai et al., 2005). Other studies have used electrophysiological measures in conjunction with backward masking tasks to understand temporal and memory processing in dyslexic adults and children (Kujala, Belitz, Tervaniemi, & Näätänen, 2003; Marler et al., 2002). Kujala et al. (2003) found impaired cortical discrimination (mismatch negativity, MMN) of sound-order reversals in dyslexic adults when an additional sound followed the tone pairs, but not when the sound preceded the pairs. The MMN is a physiological reflection of stimulus change, thought to be associated with echoic memory. Kujala et al. suggested that backward masking interferes with sound perception at the level of the cortex at a preattentive level. Marler et al. (2002) showed MMN latency delays and amplitude decreases in children with high backward masking thresholds. Additionally, they saw correlations between MMN and behavioral memory, and suggested impaired early memory mechanisms for complex sounds (Marler et al., 2002). They further hypothesized that such deficits could impair the processing of acoustic signals composed of brief spectral and temporal components.

Conclusions

These findings indicate that learning-impaired children with perceptual timing deficits show neural encoding deficits that are specific to certain, but not all, acoustic aspects of speech. Filter cues that convey the verbal message (consonant and vowel identity) are imprecisely encoded in the brainstems of these children, as evidenced by abnormal latencies of waves A, C, and O, as well as the spectral encoding of portions of F_1. At the same time, acoustic cues conveying source information related to speaker identity and prosody are encoded normally, as reflected by the latencies of waves D, E, and F, and the spectral encoding of F_0. Taken together, these findings suggest that perceptual timing deficits do not have pan-auditory neural consequences, but rather appear to emerge from a unique and selective disruption of the biological encoding of sound. The importance of these relationships and the relative ease with which speech-evoked brainstem response may be measured has led to its translation as a clinical tool—BioMAP (Biological Marker of Auditory Processing, Bio-logic, Mundelein, IL), for the assessment of sound encoding in children with learning disabilities.

Acknowledgments

We thank the children and their families for participating in this study. We would also like to thank Dr. Beverly Wright for her help with establishing our backward masking paradigm. This work was supported by National Institutes of Health Grant R01DC01510.

Reprint requests should be sent to Krista L. Johnson, Auditory Neuroscience Laboratory, Northwestern University, Frances Searle Building, 2240 Campus Drive, Evanston, IL 60208, or via e-mail: kljohnson@northwestern.edu.

Note

1. A psychoeducational test battery given to all participants included subtests taken from Woodcock–Johnson Revised (Woodcock & Johnson, 1989). These subtests were Auditory Processing (Incomplete Words and Sound Blending), Listening Comprehension, Memory for Words, Cross Out, and Word Attack. The Brief Cognitive Scale was taken from the Woodcock–Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977). Additionally, reading and spelling were assessed by using subtests from Wide Range Achievement Test-3 (Wilkinson, 1993) and phonological skills were assessed by using Comprehensive Test of Phonological Processing (Wagner & Rashotte, 1999).

REFERENCES

Abrams, D. A., Nicol, T., Zecker, S. G., & Kraus, N. (2006). Auditory brainstem timing predicts cerebral asymmetry for speech. *Journal of Neuroscience, 26,* 11131–11137.

Banai, K., Nicol, T., Zecker, S., & Kraus, N. (2005). Brainstem timing: Implications for cortical processing and literacy. *Journal of Neuroscience, 25,* 9850–9857.
Blackburn, C. C., & Sachs, M. B. (1989). Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. *Journal of Neurophysiology, 62,* 1303–1329.
Brinkmann, R. D., & Scherg, M. (1979). Human auditory on- and off-potentials of the brainstem. *Scandinavian Audiology, 8,* 27–32.
Casseday, J. H., Ehrlich, D., & Covey, E. (1994). Neural tuning for sound duration: Role of inhibitory mechanisms in the inferior colliculus. *Science, 264,* 847–850.
Casseday, J. H., Ehrlich, D., & Covey, E. (2000). Neural measurement of sound duration: Control by excitatory–inhibitory interactions in the inferior colliculus. *Journal of Neurophysiology, 84,* 1475–1487.
Cunningham, J., Nicol, T. G., Zecker, S. G., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology, 112,* 758–767.
Dobie, R. A., & Wilson, M. J. (1984). Short-latency auditory responses obtained by cross correlation. *Journal of the Acoustical Society of America, 76,* 1411–1421.
Ehrlich, D., Casseday, J. H., & Covey, E. (1997). Neural tuning to sound duration in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. *Journal of Neurophysiology, 77,* 2360–2372.
Elliott, L. L., & Hammer, M. A. (1993). Fine-grained auditory discrimination: Factor structures. *Journal of Speech & Hearing Research, 36,* 396–409.
Elliott, L. L., Hammer, M. A., & Scholl, M. E. (1989). Fine-grained auditory discrimination in normal children and children with language-learning problems. *Journal of Speech & Hearing Research, 32,* 112–119.
Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
Faure, P. A., Fremouw, T., Casseday, J. H., & Covey, E. (2003). Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. *Journal of Neuroscience, 23,* 3052–3065.
Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hearing Research, 158,* 1–27.
Frisina, R. D., Smith, R. L., & Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hearing Research, 44,* 99–122.
Gorga, M., Abbas, P., & Worthington, D. (1985). Stimulus calibration in ABR measurements. In J. Jacobsen (Ed.), *The auditory brainstem response* (pp. 49–62). San Diego, CA: College Hill Press.
Green, D. M. (1990). Stimulus selection in adaptive psychophysical procedures. *Journal of the Acoustical Society of America, 87,* 2662–2674.
Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, & Hearing Research, 43,* 1402–1415.
Hood, L. J. (1998). *Clinical applications of the auditory brainstem response.* San Diego, CA: Singular.
Inagaki, M., Tomita, Y., Takashima, S., Ohtani, K., Andoh, G., & Takeshita, K. (1987). Functional and morphometrical maturation of the brainstem auditory pathway. *Brain Development, 9,* 597–601.
Jacobson, J. T. (1985). *The auditory brainstem response.* San Diego, CA: College Hill Press.
Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). The brainstem response to speech: A biological marker of auditory processing. *Ear & Hearing, 26,* 424–434.
Johnson, K. L., Nicol, T. G., & Kraus, N. (2006). Maturation of the speech-evoked auditory brainstem response [Abstract]. *American Auditory Society.*
Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews, 84,* 541–577.
King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters, 319,* 111–115.
Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971–995.
Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science, 273,* 971–973.
Kraus, N., & Nicol, T. G. (2005). Brainstem origins for cortical ''what'' and ''where'' pathways in the auditory system. *Trends in Neurosciences, 28,* 176–181.
Kujala, T., Belitz, S., Tervaniemi, M., & Näätänen, R. (2003). Auditory sensory memory disorder in dyslexic adults as indexed by the mismatch negativity. *European Journal of Neuroscience, 17,* 1323–1327.
Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neurophysiology, 60,* 1799–1822.
Marler, J. A., & Champlin, C. A. (2005). Sensory processing of backward-masking signals in children with language-learning impairment as assessed with the auditory brainstem response. *Journal of Speech Language & Hearing Research, 48,* 189–203.
Marler, J. A., Champlin, C. A., & Gillam, R. B. (2002). Auditory memory for backward masking signals in children with language impairment. *Psychophysiology, 39,* 767–780.
Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271,* 77–81.
Møller, A. R., & Jannetta, P. J. (1983). Interpretation of brainstem auditory evoked potentials: Results from intracranial recordings in humans. *Scandinavian Audiology, 12,* 125–133.
Russo, N., Nicol, T. G., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology, 115,* 2021–2030.
Salamy, A. (1984). Maturation of the auditory brainstem response from birth through early childhood. *Journal of Clinical Neurophysiology, 1,* 293–329.
Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine, 338,* 307–312.
Sherbenou, B. L., & Johnson, S. (1997). *Test of Nonverbal Intelligence-3.* Austin, TX: Pro-Ed.
Sininger, Y., & Starr, A. (Eds.) (2001). *Auditory neuropathy: A new perspective on hearing disorders.* Albany, NY: Singular Thomson Learning.
Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography & Clinical Neurophysiology, 39,* 465–472.
Sohmer, H., Pratt, H., & Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography & Clinical Neurophysiology, 42,* 656–664.
Stevens, K., & Blumstein, S. (1978). Invariant cues for place

of articulation in stop consonants. *Journal of the Acoustical Society of America, 64,* 1358–1368.
Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9,* 182–198.
Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271,* 81–84.
Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences, 682,* 27–47.
Tallal, P., Stark, R. E., & Mellits, E. D. (1985). Identification of language-impaired children on the basis of rapid perception and production skills. *Brain & Language, 25,* 314–322.
Temple, E., Poldrack, R. A., Protopapas, A., Nagarajan, S., Salz, T., Tallal, P., et al. (2000). Disruption of the neural response to rapid acoustic stimuli in dyslexia: Evidence from functional MRI. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 13907–13912.
Torgeson, J. K. (1991). Learning disabilities: Historical and conceptual issues. In B. Y. L. Wong (Ed.), *Learning about learning disabilities.* San Diego, CA: Academic Press.
Wagner, R. J. K. T., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing (CTOPP).* Austin, TX: PRO-ED.
Wang, X., & Sachs, M. B. (1994). Neural encoding of single-formant stimuli in the cat. II. Responses of anteroventral cochlear nucleus units. *Journal of Neurophysiology, 71,* 59–78.
Watanabe, T. (1978). Responses of the cat's collicular auditory neuron to human speech. *Journal of the Acoustical Society of America, 64,* 333–337.
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. *Biological Psychology, 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.
Wilkinson, G. (1993). *Wide Range Achievement Test-3.* Wilmington, DE: Jastak & Associates.
Wilson, M. J., & Dobie, R. A. (1987). Human short-latency auditory responses obtained by cross-correlation. *Electroencephalography & Clinical Neurophysiology, 66,* 529–538.
Woodcock, R., & Johnson, M. (1977). *Woodcock–Johnson Psycho-Educational Battery. Tests of cognitive ability.* Allen, TX: DLM Teaching Resources.
Woodcock, R., & Johnson, M. (1989). *Woodcock–Johnson Psycho-Educational Battery—Revised. Tests of cognitive ability.* Allen, TX: DLM Teaching Resources.
Worden, F. G., & Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography & Clinical Neurophysiology, 25,* 42–52.
Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387,* 176–178.