Audiology&
Neurotology

Subcortical Laterality of Speech Encoding

Jane Hornickel[a]   Erika Skoe[a]   Nina Kraus[a–c]

[a] Auditory Neuroscience Lab., Roxelyn and Richard Pepper Department of Communication Sciences and Disorders,
[b] Department of Neurobiology and Physiology, and [c] Department of Otolaryngology, Northwestern University,
Evanston, Ill., USA

Abstract

It is well established that in the majority of the population language processing is lateralized to the left hemisphere. Evidence suggests that lateralization is also present in the brainstem. In the current study, the syllable /da/ was presented monaurally to the right and left ears and electrophysiological responses from the brainstem were recorded in adults with symmetrical interaural click-evoked responses. Responses to the right-ear presentation occurred earlier than those to left-ear presentation in two peaks of the frequency following response (FFR) and approached significance for the third peak of the FFR and the offset peak. Interestingly, there were no differences in interpeak latencies indicating the response to right-ear presentation simply occurred earlier over this region. Analyses also showed more robust frequency encoding when stimuli were presented to the right ear than the left ear. The effect was found for the harmonics of the fundamental that correspond to the first formant of the stimulus, but was not seen in the fundamental frequency range. The results suggest that left lateralization of processing acoustic elements important for discriminating speech extends to the auditory brainstem and that these effects are speech specific.

Copyright © 2008 S. Karger AG, Basel

Introduction

Cortical asymmetry of language processing, as determined by functional imaging, electrophysiological responses, and performance on dichotic listening tasks, is well established. The left hemisphere appears to be specialized for processing language, and this specialization may be partly due to the acoustic characteristics of speech, dominated by transient elements and fast temporal transitions. Music has more sustained temporal and spectral elements, slower transitions, and finer frequency spacing than speech, and is primarily processed in the right hemisphere. A prominent theory of lateralization of speech and music processing suggests that the left hemisphere has maximum temporal resolution at the expense of frequency resolution while the right hemisphere shows the opposite pattern [Zatorre et al., 2002]. Imaging studies have supported this theory, showing greater activation in the left auditory cortex than the right to rapid temporal changes in stimuli [Zatorre and Belin, 2001; Schonwiesner et al., 2005]. Rapid spectral changes may also be processed by the left hemisphere. A recent study found greater left hemisphere auditory cortex activity to stimuli with frequency sweeps that were 25–50 ms long while the right hemisphere showed greater activity to the same stimuli when 200–300 ms long [Boemio et al., 2005]. Abrams et al. [2008] found a similar pattern for speech stimuli in that the right auditory cortex had larger and more accurate representations of the slow temporal envelope of the speech

signal, which corresponds to the syllable rate pattern in speech. Thus it appears that certain temporal aspects of acoustic signals are lateralized to each hemisphere.

In addition to the functional imaging evidence for left lateralization of speech acoustics processing, hemispheric specialization is also predicted from behavioral accounts of a right-ear advantage (REA) during dichotic listening tasks with speech stimuli. The commonly reported REA for speech processing results in faster reaction times, greater accuracy, and clearer perception of speech when presented to the right than the left ear [Sidtis, 1982; Schwartz and Tallal, 1980; Shtyrov et al., 2000; Spellacy and Blumstein, 1970]. Similarly, a left-ear advantage has been found for pitch discrimination of tone stimuli [Sidtis, 1982]. In accordance with the aforementioned neural data, the REA for speech appears to be related to processing of rapid acoustic changes. When the formant transitions of stop consonant CV (consonant-vowel) syllables were lengthened from 40 to 80 ms, the magnitude of the REA for speech decreased [Schwartz and Tallal, 1980].

Electrophysiological correlates of the REA for speech have been found with cortical evoked potentials. Participants determined behaviorally to have an REA have earlier N100 latency over the left temporal lobe and greater P300 amplitude in the left hemisphere during dichotic listening studies with speech stimuli [Ahonniska et al., 1993; Eichele et al., 2005]. Similarly, the mismatch negativity (which reflects acoustic changes in a stimulus sequence) elicited by speech stimuli was more left lateralized in participants displaying an REA compared to those who did not display an REA [Mathiak et al., 2000]. Relationships between cortical laterality and subcortical function have recently been investigated within the same subject. Children with delayed latencies in evoked brainstem responses to speech stimuli showed reduced cortical amplitude asymmetries in response to speech stimuli relative to children with earlier latencies [Abrams et al., 2006]. This effect was especially prevalent over the temporal electrodes and a calculated asymmetry index was found to be correlated with subcortical onset latency such that earlier onset latencies were related to greater leftward asymmetries in amplitude [Abrams et al., 2006]. These findings suggest that the subcortical responses and cortical lateralization are interrelated.

Differences in the peripheral auditory system also support the right-ear/left-hemisphere specialization for rapid temporal signal processing. Spontaneous otoacoustic emissions were found to be more prevalent in the right ear than the left and transient evoked otoacoustic emissions (TEOAEs) in the right ear had greater amplitudes and signal-to-noise ratios than those in the left [Morlet, 1995; Driscoll et al., 1999, 2002]. When comparing TEOAEs to distortion product otoacoustic emissions (DPOAEs) in infants, Sininger and Cone-Wesson [2004] found that TEOAEs were more robust in the right ear while DPOAEs were more robust in the left, indicating that the right ear appeared to preferentially amplify temporal acoustic features while the left preferentially amplified more tonal features. Thus it appears that the structure and physiology of the right ear could prime the system for left hemispheric lateralization of speech processing.

The lateralization of transient signal processing is also evident subcortically in the brainstem, as evidenced by click-evoked auditory brainstem responses (ABRs) [Levine et al., 1988; Sininger et al., 1998; Sininger and Cone-Wesson, 2006]. Infants showed interaural latency and amplitude differences in responses to clicks; responses to right-ear stimulation were earlier and larger than responses to left-ear presentation [Sininger et al., 1998; Sininger and Cone-Wesson, 2006]. Interaural ABR amplitude (but not latency) differences in response to clicks have also been documented in adults [Decker and Howe, 1981; Levine et al., 1988; Spivak and Seitz, 1988]. One study found wave V interaural amplitude differences to be positively correlated with wave V latency differences [Spivak and Seitz, 1988]. The infant studies described above had nearly 1000 participants, while the adult studies had approximately 50; therefore, the absence of latency differences in adults may simply be due to the extreme subtlety of the effect.

Additional studies have shown subcortical lateralization of spectral encoding. Tonal (and other highly periodic) stimuli elicit a frequency following response (FFR) in which the periodicity of the response peaks matches the periodicity of the stimulus. The FFR is thought to be generated by a series of brainstem nuclei, including the inferior colliculus and lateral lemniscus, and represents the temporal coding of frequency through neural phase-locking [Smith et al., 1975; Hoormann et al., 1992]. Previous work on adults found that responses to left- and right-ear presentation of tone bursts differed significantly in normalized amplitude of the FFR region [Ballachanda et al., 1994, 2000]. In comparing the right- and left-ear contributions to the binaural response, the magnitude of the binaural response was more greatly attenuated when subtracting out the left than the right response, suggesting the left ear contributed more to the binaural response [Ballachanda et al., 1994]. In summary, prior evidence suggests lateralization of simple stimuli at the level of the

brainstem. Transient, broadband clicks elicit earlier and larger responses when presented to the right ear. In contrast, tone burst stimuli elicit larger responses when presented to the left ear. This dichotomy suggests selective lateralization of brainstem auditory processing.

While previous studies of subcortical lateralization employed click or tone burst stimuli, the current study assessed subcortical encoding of speech stimuli in order to examine whether a right-ear/left-hemisphere lateralization for speech processing extends to the auditory brainstem. The /da/ stimulus used in the current study has rich temporal and spectral characteristics and is acoustically similar to a click followed by a tone. Unlike tones, however, the stimulus contains rapid but broad spectral changes. Due to these acoustic similarities, responses to the /da/ syllable include regions similar to the click ABR and tone FFR. Brainstem responses to CV syllables are characterized by transient and sustained elements that mimic the acoustic signal of the stimulus with considerable fidelity [Kraus and Nicol, 2005, see fig. 1], to the degree that the stimulus is intelligible when the neural response is played back as a sound [Galbraith et al., 1995]. The transient elements and harmonic content of the response represent the filter characteristics of speech (timing and spectral elements), corresponding to the formants which are important for distinguishing phonemes (i.e. distinguishing /da/ from /ga/). The fundamental frequency, determined by glottal pulsing, gives rise to the perception of pitch and is important for prosody. In the speech-evoked brainstem response, the fundamental frequency (F_0) is apparent in the spectral energy corresponding to the F_0 (frequency domain) and in the major peaks of the FFR, which occur at a time interval corresponding to the period of the F_0 (time domain).

The current study investigated subcortical asymmetry of speech encoding in normal-hearing adults by recording brainstem responses to a 40-ms syllable, /da/, presented independently to the right and left ears. This stimulus simultaneously contains broad spectral and fast temporal information characteristics of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. Responses from presentations to the right and left ears were analyzed in the time and frequency domains and were compared to each other as well as to the stimulus. Based on prior research suggesting an REA for transient, speech-like stimuli, we hypothesized that the right-ear responses would have reduced onset and offset latencies relative to the left. Because of the rapid spectral changes in the stimulus, presentation to the right ear was also hypothesized to elicit stronger spectral encoding (in the formant range) than presentation to the left ear. For fundamental frequency encoding, no differences or the opposite pattern were predicted, as this element represents a pitch cue.

Methods

Participants

Twelve adults (9 female), aged 21–30 years (mean = 25.67), participated in the study. All were right-handed by self-report and as determined by the Edinburgh Handedness Inventory [Oldfield, 1971; 11 participants: mean = 85, range = 54–100]. Air conduction audiograms were measured upon arrival to determine symmetrical hearing. All participants had pure tone thresholds ≤20 dB SPL for octave frequencies 250–8000 Hz. Pure tone averages were calculated from 500-, 1000-, and 2000-Hz thresholds (PTA1) and from 1000-, 2000-, and 4000-Hz (PTA2) thresholds. PTA1 and PTA2 did not differ significantly between the ears (interaural difference for PTA1: mean = 1.97 dB, range = 0–5; PTA2: mean = 2.12 dB, range = 0–5).

Procedure

After the initial hearing screening, participants watched a movie of their choice with subtitles (soundtrack muted) while neurophysiological data were collected. Electrophysiological testing lasted approximately 1 h. To obtain reliable brainstem responses, with as little movement artifact as possible, participants needed to stay still and relaxed for the duration of testing. Our laboratory has found that subjects are more compliant when they are allowed to watch a movie during the recording sessions. Participants were monetarily compensated for their participation. All procedures were approved by the internal review board of Northwestern University and informed consent was obtained from all participants.

Stimuli and Recording Parameters

Replications of auditory brainstem responses to clicks were collected before and after presentation of the speech stimulus. The click stimulus was a broadband spectral square wave lasting 100 μs. The /da/ stimulus was a 40-ms synthesized speech syllable produced in KLATT [Klatt, 1980] with a fundamental frequency (F_0) that rose linearly from 103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 ms. The first formant (F_1) rose from 220 to 720 Hz, while the second formant (F_2) decreased from 1700 to 1240 Hz over the duration of the stimulus. The third formant (F_3) fell slightly from 2580 to 2500 Hz, while the fourth (F_4) and fifth formants (F_5) remained constant at 3600 and 4500 Hz, respectively. The stimulus was comprised of the initial noise burst and the transition of formant frequencies between the consonant and a steady-state vowel. Although the steady-state portion was not present, the stimulus was still perceived as being a CV syllable.

Stimuli were presented monaurally through insert earphones (ER-3; Etymotic Research, Elk Grove Village, Ill., USA) at 80.3 dB SPL. Click stimuli were presented at 13.3 Hz and speech stimuli were presented in alternating polarities at 10.9 Hz (interstimulus interval = 51 ms). Stimulus presentation was in a block design; in

Fig. 1. Time domain waveforms of the stimulus (**a**) and the grand average brainstem response for the right-ear presentation (**b**). Peaks D, E, F, and O of the response mimic the large troughs of the stimulus. Stimulus is time-shifted 8 ms in order to facilitate its visual comparison with the response by accounting for the neural conduction lag.
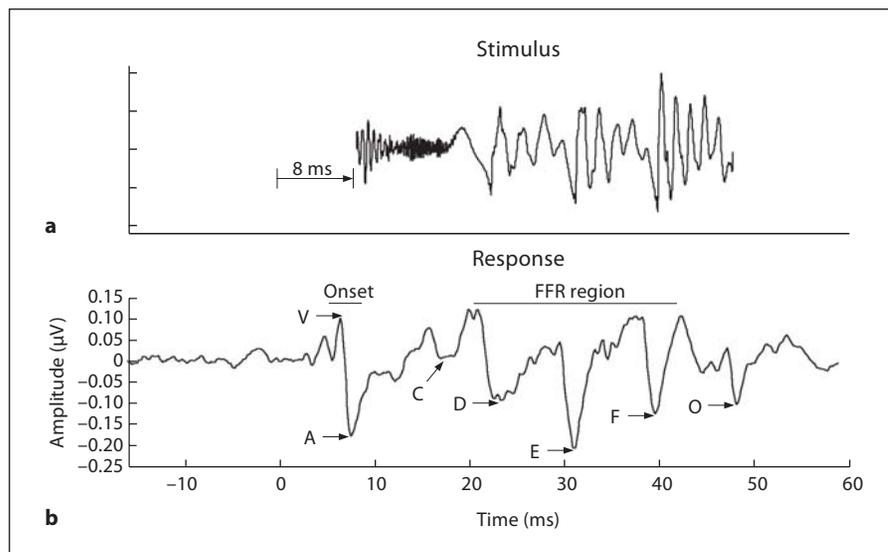
each block, sounds were delivered only to a single ear (i.e. stimuli were not interleaved between the two ears within a block). Across subjects, the order of the blocks was randomized, and the two transducers were swapped (similar to Ballachanda et al. [1994]). Both ears were occluded with insert earphones throughout the session, regardless of which ear was stimulated.

A vertical montage of three Ag-AgCl electrodes was used to record neurophysiological responses (central vertex, forehead ground, and ipsilateral earlobe reference). Responses to both click and speech stimuli were recorded with the Bio-logic system (Bio-MAP software, NavigatorPro AEP system, Bio-logic Systems Corp., a Natus Company, Mundelein, Ill., USA). To minimize cochlear microphonic and stimulus artifact, responses to alternating polarities of the /da/ stimulus were added together. Artifact rejection was executed online with a criterion of ± 23 μV. Three blocks of 2000 accepted sweeps were collected for each ear and averaged using a 74.67-ms time window that included 15.8 ms of prestimulus and 58.87 ms of postonset activity. The responses were online bandpass filtered from 100 to 2000 Hz (12 db/octave) and digitally sampled at 6857 Hz. 6857 Hz (approx. 0.146 ms/point) was the maximum sampling rate permitted by the Bio-logic software given the specified time window (512 points/74.67 ms) and this allowed for a frequency resolution up to 3428.5 Hz. Responses to click stimuli were digitized at a 24.2-kHz sampling rate over a 10.66-ms time window.

Data Processing: Speech- and Click-Evoked Brainstem Response

Data analysis followed published reports using similar stimulus and recording parameters (see Russo et al. [2004] for an indepth review of these procedures).

The characteristic seven peaks of the response to /da/ (V, A, C, D, E, F, O) were manually identified by the experimenter and confirmed by an experienced observer, both blind to the ear of presentation. The onset burst (fig. 1) contains broad frequency information and elicits wave V, as well as a trough following V, known as wave A. Peak C is thought to encode the transition from the

aperiodic stop burst to the periodic (voiced) formant transition, while peak O corresponds to the cessation of the stimulus. The FFR includes peaks D, E, and F, which occur at a period approximating the fundamental frequency of the stimulus and correspond to the voiced portion of the syllable. Higher frequency information, such as formants, is encoded in the smaller voltage fluctuations between the three FFR waves. Peaks that were deemed not replicable or not reliably above the noise floor were marked as missing data points (see below). The same observers also marked peak V of the click-evoked response.

All data analysis was automated using routines coded in Matlab 7 (The MathWorks, Inc., Natick, Mass., USA). To obtain measures of peak latency, the local minima (the maxima, in the case of wave V) within two sampling points (± 2) of the chosen peak were identified. For wave V, a narrower range was used (+2) to avoid the accidental identification of wave IV.

Frequency encoding was analyzed using a Fourier analysis of two different time windows that include peaks D, E, and F of the response (11.4–40.6 ms and 21.9–40.6 ms). The two windows differ on their inclusion of peak C, which was deemed unreliable in the present study as it was reliably above the noise floor in only 66% of the participants and was consequently excluded from latency analyses. To increase the number of sampling points in the frequency domain, the time window was zero-padded to 4096 points before performing a discrete Fourier transform. Average spectral amplitude was calculated for three frequency ranges: fundamental frequency (F_0) 103–120 Hz, first formant (F_1) 455–720 Hz, and high frequency (HF) 721–1154 Hz. The first formant of the stimulus ramps from 220 to 720 Hz over the 40-ms syllable. The F_1 frequency range used for FFR analysis accounts for the time lag and the corresponding F_1 frequency ramping between the onset of the stimulus and the periodic formant transition that elicits the FFR. The HF range corresponded to the 7th through 11th harmonics of the F_0 of the stimulus, a frequency range between the first and second formants. The second formant is beyond the phase-locking capabilities of the brainstem response; therefore F_2–F_5 were not included in the analysis [Liu et al., 2006].

Correlations between the stimulus and response (SR) and right- and left-ear (RL) responses were calculated with a cross-correlational technique. The cross-correlational method is a robust measure of latency differences because it is objective and not dependent on manual peak identification. This technique measures the direction and linear relationship between two signals. One signal is shifted in time with respect to the other (stationary) signal to find the time lag for which the morphology of the two signals is most similar, with larger Pearson's correlation coefficients (r) indicating greater coherence. SR correlations measure the degree to which the response mimics the stimulus, and RL response correlations identify the effect of ear on the latency and morphology of the response. Correlations were performed on a low-pass-filtered version of the stimulus (2 kHz). Low-pass filtering removes the consonant noise burst and allows for higher correlations given the low-pass filter nature of the brainstem response [Liu et al., 2006]. The response (20–40 ms) was correlated with the sustained formant transition portion of the stimulus (13–34 ms). To account for the neural conduction delay between the onset of the stimulus and the onset of the response (approx. 7–10 ms), the strongest correlation was found between the filtered stimulus and a 6.6- to 9.9-ms time-shifted version of the response. The maximum correlation coefficient and the corresponding lag were calculated separately for the left-ear presentation and the right-ear presentation. RL correlations were calculated using the same method as the SR correlations. The left-ear response was shifted temporally over the right-ear response within a ±1.5-ms lag range (allowing for the possibility that the left response occurred before the right). The straight correlation (zero lag) was measured in addition to the maximum correlation and its respective time lag. No difference between the zero-lag r and the maximum-obtainable r would suggest that there was no difference in ear of presentation using this measure. However, a significant difference would indicate that responses to one ear were consistently lagging behind (i.e. later) compared to the responses to the other. Given the nonnormal distribution of Pearson's correlation coefficients, r values were converted to z-scores using Fisher transform before performing statistical analyses.

Measures obtained from time domain, frequency domain and correlational analyses were analyzed using two-tailed paired t tests (right-ear presentation vs. left-ear presentation) conducted in SPSS (SPSS Inc., Chicago, Ill., USA). Missing data were replaced with the overall mean for that measure in the respective ear. Effect sizes, which indicate the strength of the effect independent of sample size, were calculated using Cohen's d.

Results

Click Responses
Wave V latencies were not significantly different for the right- and left-ear presentations [t(11) = 0.193, n.s.]. All wave V latencies and interaural latency differences were within clinical norms for adults as reported by Hall [1992] and were consistent with norms established in-house. As a group there were no tendencies toward consistent interaural latency differences, with 41.6% of the

participants showing shorter wave V latencies for right-ear presentation, 41.6% for left-ear presentation, and 16.6% showing no interaural latency differences. Wave V amplitudes also did not differ either between the left and right-ear presentations [right: mean = 0.127 μV, SD = 0.04; left: mean = 0.137 μV, SD = 0.04; t(11) = n.s.; table 1].

Measures of Timing
Peak Latency
Latencies of all major peaks of the speech response were analyzed (table 1). Peak C was excluded from the analysis due to its poor detectability (present in only two thirds of the responses). The onset peaks (V and A) were 100% detectable in all subjects and conditions, and the remaining peaks were also very robust (table 1). The latencies of peaks D and F were found to differ significantly between the left- and right-ear presentations with D and F having earlier latencies in the response to right-ear presentation compared to the left [D: t(11) = −2.747, p < 0.05, d = 0.303; F: t(11) = −3.226, p < 0.01, d = 0.750; fig. 2]. Peak E and O latency differences approached the significance threshold (p = 0.073 and p = 0.058, respectively), while differences were not found for the onset peak latencies (V and A). Interpeak latencies of peaks D, E, and F, which reflect the period of the fundamental frequency of the stimulus, were not significantly different between the ears [D–E: t(11) = −0.932, n.s.; E–F: t(11) = 0.162, n.s.; table 1].

SR Correlations
The left- and right-ear presentations resulted in responses with very similar morphology (fig. 3) and the maximum SR correlation did not differ between the two [left: r = 0.277, right: r = 0.290; t(11) = 0.708, n.s.]. Despite a lack of difference in strength of SR correlation, interaural timing differences were detected using this technique. Based on differences in the time shift needed to obtain the highest correlation, the responses to left-ear presentation were found to be delayed with respect to the right. Shorter time shifts were needed to maximize the SR correlation with responses to right-ear presentation compared to the left [right: mean = 8.376, SD = 0.59; left: mean = 8.765, SD = 0.68; t(11) = −2.886, p < 0.05, d = 0.610].

RL Correlations
The results from the RL correlations lend further support for the timing delays observed in the peak latency and SR correlation analysis. A maximum RL correlation
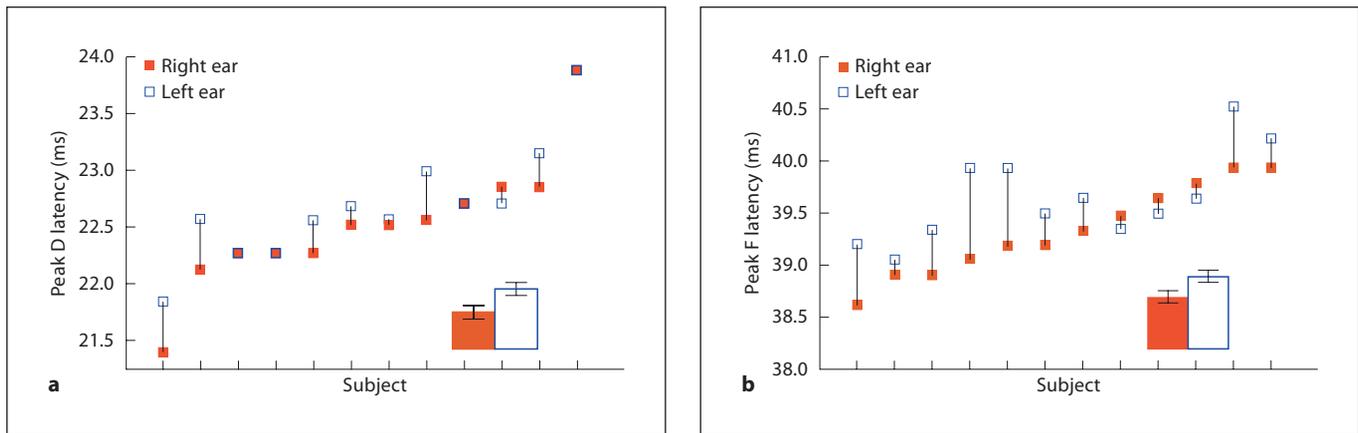
Fig. 2. Dot plots of interaural latency differences for each participant for right- and left-ear presentation for peak D (**a**) and F (**b**). Bar graphs of the mean latencies for each peak are included in the insets (D: right ear 22.52 ms, left ear 22.68 ms; F: right ear 39.33 ms, left ear 39.65 ms). Standard errors are plotted in the insets.

Table 1. Means (standard deviations) of peak latency, correlational, and spectral encoding measures

Measure	Right ear	Left ear	t	p	Effect size	% Detectability
Latency measures, ms						
Click V	5.51 (0.26)	5.50 (0.26)	0.19	0.850		100
V	6.64 (0.27)	6.58 (0.25)	0.90	0.389		100
A	7.65 (0.38)	7.61 (0.33)	0.71	0.495		100
D	22.52 (0.58)	22.68 (0.51)	−2.75	0.019*	0.30	87.5
E	30.96 (0.38)	31.28 (0.58)	−1.94	0.078	0.65	91.6
F	39.33 (0.43)	39.65 (0.43)	−3.23	0.008**	0.75	91.6
O	48.14 (0.39)	48.37 (0.58)	−2.12	0.058	0.47	83.3
D–E	8.41 (0.48)	8.58 (0.40)	−0.93	0.371		
E–F	8.41 (0.43)	8.38 (0.42)	0.16	0.874		
Correlation measures						
SR (20–40 ms)	0.29 (0.05)	0.28 (0.09)	0.71	0.494		
SR lag	8.38 (0.59)	8.77 (0.68)	−2.89	0.015*	0.61	
Spectral amplitude measures						
F_0 (21.9–40.6 ms)	6.03 (3.45)	5.39 (2.39)	0.92	0.377		
F_1 (21.9–40.6 ms)	0.80 (0.29)	0.65 (0.26)	3.2	0.008**	0.51	
HF (21.9–40.6 ms)	0.36 (0.11)	0.32 (0.09)	2.71	0.020*	0.41	
F_0 (11.4–40.6 ms)	6.97 (5.17)	6.00 (2.73)	0.73	0.484		
F_1 (11.4–40.6 ms)	0.91 (0.33)	0.78 (0.28)	3.62	0.004**	0.45	
HF (11.4–40.6 ms)	0.40 (0.12)	0.34 (0.10)	3.34	0.007**	0.50	

* $p < 0.05$; ** $p < 0.01$.

was obtained when the response to left-ear presentation was shifted later in time by an average of 0.219 ms (SD = 0.28). This maximum correlation was significantly greater than the straight correlation (zero lag) [r = 0.806 and 0.789, respectively; t(11) = 2.828, p < 0.05, d = 0.135].

Measures of Spectral Encoding
Fourier Analyses

Relative to the left ear, the right-ear responses showed increased spectral amplitudes for frequencies above F_0. Specifically, the right-ear presentation evoked greater

Fig. 3. Time domain waveform of the stimulus (13–34 ms; **a**) and right-ear (red line) and left-ear (blue line) responses of a single participant over the FFR region (**b**). Interaural latency differences (ID) are given for peaks D and F. Peak E showed no interaural differences for this participant. The three corresponding peaks in the stimulus are marked with small vertical lines.
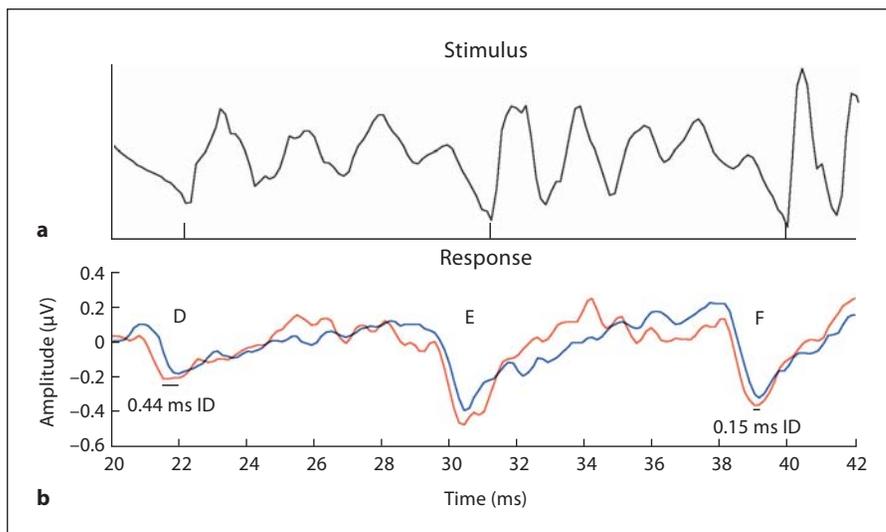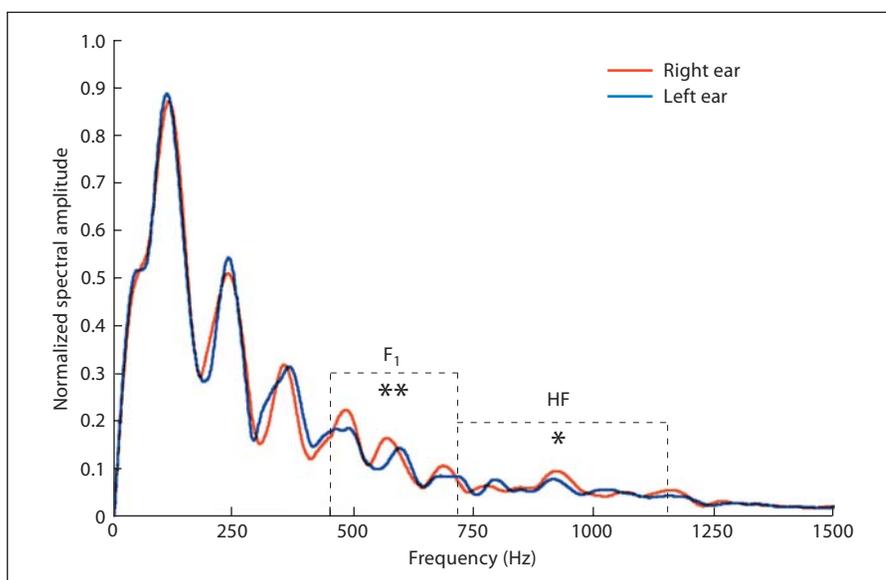

Fig. 4. Grand average normalized spectral amplitude for right- and left-ear responses over the 21.9- to 40.6-ms time range encompassing the FFR. For graphing purposes, individual spectra were normalized to the spectral maxima (approx. 100 Hz) for each individual's response and then the resulting normalized spectra were averaged together. $* p < 0.05$; $** p < 0.01$.

spectral amplitude for both the F_1 and HF ranges than the left. This result was consistent for both FFR time ranges (table 1; fig. 4). In contrast, there were no significant interaural differences in the encoding of the fundamental frequency in the longer and shorter FFR time windows [$t(11) = 0.725$, n.s.; $t(11) = 0.921$, n.s., respectively]. Symmetrical encoding of the fundamental frequency was further supported by the overall similarity in the waveform morphology, concordant SR correlation coefficients, and lack of interpeak FFR latency differences.

Discussion

The current study sought to identify lateralization of speech encoding at the level of the brainstem by presenting CV stimuli to the right and left ears. Results confirm an REA for specific acoustic features that are characteristic of speech. These effects were exclusive to the speech stimulus as there were no interaural differences in click-evoked responses. Previous work has shown that musical experience and language experience enhance brainstem encoding of speech sounds [Krishnan et al., 2005; Musacchia et al., 2007], and the results of the current study are consistent with the notion that language exposure can

shape lower-level acoustic processing. In Krishnan et al. [2005], native Mandarin speakers were found to have stronger tracking of Mandarin tones (variable pitch speech stimuli) than native English speakers. Johnson et al. [2008] also found that click- and speech-evoked brainstem responses have different developmental trajectories consistent with experience-dependent shaping of auditory function (different exposure to and use of spoken language). Musical experience has also been shown to influence the auditory brainstem response to speech. Musacchia et al. [2007] and Wong et al. [2007] found that musicians had enhanced responses to speech stimuli compared to nonmusicians. While statements of causality cannot be made, these results suggest that extended and consistent auditory stimulation can lead to stronger encoding of speech sounds at the level of the brainstem, even when the auditory training does not explicitly involve speech sounds. In the current study, it was the more ecologically valid speech stimulus and not the click stimulus that evoked differences in the brainstem response to the right-ear versus left-ear stimulation. It is possible that the contralateral projections from the right ear to the left hemisphere have been reinforced and subsequently enhanced by the exposure and everyday use of speech; however, the current study did not directly compare responses to speech and nonspeech stimuli with similar acoustic features (e.g. speech played backwards that has no linguistic meaning but the same acoustics) and the influence of language experience on the observed effects is still speculative.

Measures of Timing
Latency Analyses

Given what is known about the REA for speech and how it is linked to the processing of transient, speech-like stimuli in the left hemisphere, interaural latency differences were predicted in the speech-evoked ABR [Schwartz and Tallal, 1980; Zatorre and Belin, 2001; Schonwiesner et al., 2005]. In the current study, latency differences were found for peaks D and F and approached significance for peaks E and O. In all cases, including peaks E and O, the responses to right-ear presentation occurred earlier than those to left-ear presentation. Despite the interaural differences in absolute peak latency, there was no difference between the responses in the interpeak latencies within the FFR peaks, indicating both responses had the same periodicity, such that response elements reflecting the F_0 were symmetrical.

There were, however, no latency differences for the onset peaks (V and A), mimicking the current click wave V results. Previous work in children has shown that latency

of peaks V and A of click-evoked and speech-evoked brainstem responses are highly correlated [Song et al., 2006]. This, in addition to infant data showing reduced peak V latency in right-ear presentation relative to left [Sininger et al., 1998; Sininger and Cone-Wesson, 2006], would suggest that the speech-evoked peak V latency would also be earlier when stimuli were presented to the right ear compared to the left, but those infant data were collected with a large subject pool and the negative findings of the present study may be due to small sample size. Furthermore, the fact that the latency differences do not encompass these early peaks and are apparent in a relatively small sample suggests that the REA may not emerge until after the initial stimulus noise burst or that the REA is simply more robust and apparent in these later peaks. FFR latency has not been found to be correlated with the characteristic peaks of the click-evoked response [Hoormann et al., 1992]. It is possible that the right-ear/left-hemisphere pathway contains a more efficient phase-locking network that results in interaural latency differences during the FFR region but not for the onset or click responses. Moreover, speech-evoked brainstem responses in children have shown that differences in the representations of the speech sounds /ba/, /da/, and /ga/ do not occur in the onset response, but emerge later in the more periodic portion of the response corresponding to the format transition [Johnson et al., in press]. The lack of differences in the onset responses in the current study and the study by Johnson et al. [in press] might also be attributable to the acoustic differences between the transient short-duration noise burst of the stop consonants (10 ms in both studies) and the more dynamic and harmonically rich, voiced formant transition.

Correlational Analyses

The peak latency differences in the FFR region were also confirmed using SR and RL cross-correlations. The right-ear presentation and left-ear presentation responses had comparable SR correlation coefficients, suggesting that the two were similar in global morphology. In order to calculate the highest correlation, the response was shifted in time relative to the stimulus. Compared to the right-ear presentation, the response to left-ear presentation required a greater time shift (lag) to obtain an equivalent SR correlation. The responses were also compared directly using RL correlations. Again, the FFR to left-ear presentation was found to lag behind the right. Along with evidence of hemispheric latency differences in cortical responses to CV stimuli [Eichele et al., 2005], the latency differences found in the present study support an

REA for transient elements of speech-like stimuli in the brainstem similar to effects that have previously been documented in the cortex [Schonwiesner et al., 2005; Zatorre and Belin, 2001].

Measures of Spectral Encoding
Fourier Analyses
Relative to the left ear, presentation of the /da/ syllable to the right ear resulted in increased amplitude of frequency encoding in frequency ranges corresponding to the first formant and harmonics of the stimulus between the first and second formants. These effects were found over the FFR and a more global sample of the response. The two responses did not differ in their encoding of the fundamental frequency. Taken together with the absence of interaural interpeak latencies between the FFR peaks, indicating both responses had the same periodicity, it appears that response elements reflecting the F_0 are symmetrically processed.

Although previous behavioral and neurophysiological results suggest a left-ear advantage for tonal (sine-wave and square-wave) stimuli [Sidtis, 1982; Ballachanda et al., 1994, 2000], the stimulus used in the present study was harmonically rich and contained rapid, broad frequency changes. The periodic portion of the speech stimulus was too short to be perceptually identified as a tone [Robinson and Patterson, 1995] and contained a rich harmonic complex unlike the stimuli used in Ballachanda et al. [1994, 2000]. These factors may account for the symmetrical rather than left-ear enhancement of encoding of the fundamental frequency.

Summary and Conclusions

The results of the current study suggest that the temporal and harmonic elements of the speech signal are preferentially encoded by the right-ear/left-hemisphere pathway, but that the fundamental frequency, perceived as pitch, is not. Kraus and Nicol [2005] have observed dissociation between the encoding of the fundamental frequency (source) and harmonic and timing cues (filter) in the brainstem. Brainstem responses to these acoustic streams have also been found to differ between normal and clinical populations, and the current results are consistent with those findings. Children with language-based learning disorders, who tend to have particular difficulty with phonemic contrasts, have delayed brainstem responses relative to their normal learning peers [King et al., 2002; Wible et al., 2004; Banai et al., 2005, 2008]. These

children also differ from their normal learning peers in harmonic encoding, but not in the spectral amplitude of the fundamental frequency [Wible et al., 2004; Banai et al., 2008], suggesting the transient and harmonic elements are important for phoneme discrimination but the pitch is not. On the other hand, children with autism spectrum disorders show reduced pitch tracking in the brainstem relative to normally developing children [Russo et al., 2008]. Unlike children with language-based learning disorders, children with autism spectrum disorders show impairments in social interactions, thought to be mediated by a deficit in pitch encoding which affects perception of prosody. The interaural differences found in this study reflect differences in timing and harmonics of the signal but not the fundamental frequency, suggesting the pathway is primed for encoding elements of the phonemic content of speech but not the prosodic elements.

To our knowledge, this is the first study to identify differences in speech-evoked brainstem responses between right- and left-ear stimulation. The results showed that the right-ear presentation of a stop consonant syllable elicited stronger formant frequency encoding than the left, including a frequency range important for distinguishing phonemes. Right-ear presentation also resulted in an earlier response over the FFR region than the left. While responses to right- and left-ear stimulation did not differ in the quality of stimulus replication (maximum SR correlation), responses to a right-ear stimulus occurred earlier, suggesting the auditory system is predisposed to respond more quickly to right-ear presentation of speech sounds. The results of the current study showed differences in responses to right- and left-ear presentation reflecting the filter elements of the speech signal (higher frequency spectrum and timing) but a selective absence of differences in the source elements (fundamental frequency). This finding suggests a lateralization of processing in the auditory brainstem for selective stimulus components and bolsters the existence of REA for speech and speech-like stimuli. Further investigation of responses to nonspeech stimuli would determine the extent to which this effect is due to stimulus complexity or to linguistic influence on subcortical processes.

References

Abrams DA, Nicol T, Zecker SG, Kraus N: Auditory brainstem timing predicts cerebral asymmetry for speech. J Neurosci 2006;26:11131–11137.

Abrams DA, Nicol T, Zecker SG, Kraus N: Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. J Neurosci 2008;28:3958–3965.

Ahonniska J, Cantell M, Tolvanen A, Lyytinen H: Speech perception and brain laterality: the effect of ear advantage on auditory event-related potentials. Brain Lang 1993;45:127–146.

Ballachanda BB, Moushegian G: Frequency-following response: effects of interaural time and intensity differences. J Am Acad Audiol 2000;11:1–11.

Ballachanda BB, Rupert A, Moushegian G: Asymmetric frequency-following responses. J Am Acad Audiol 1994;5:133–137.

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N: Reading and subcortical auditory function. 10th Int Conf Cogn Neurosci, Bodrum, 2008.

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

Boemio A, Fromm S, Braun A, Poeppel D: Hierarchical and asymmetry temporal sensitivity in human auditory cortices. Nat Neurosci 2005;8:389–395.

Decker TN, Howe SW: Auditory tract asymmetry in brainstem electrical responses during binaural stimulation. J Acoust Soc Am 1981;69:1084–1090.

Driscoll C, Kei J, McPherson B: Handedness effects on transient evoked otoacoustic emissions in schoolchildren. J Am Acad Audiol 2002;13:403–406.

Driscoll C, Kei J, Murdoch B, McPherson B, Smyth V, Latham S, et al: Transient evoked otoacoustic emissions in two-month-old infants: a normative study. Audiology 1999;38:181–186.

Eichele T, Nordby H, Rimol LM, Hugdahl K: Asymmetry of evoked potential latency to speech sounds predicts the ear advantage in dichotic listening. Brain Res Cogn Brain Res 2005;24:405–412.

Galbraith GCA, Paul W, Branski R, Comerci N, Rector PM: Intelligible speech encoded in the human brain stem frequency-following response. Neuroreport 1995;6:2363–2367.

Hall JWI: Handbook of Auditory Evoked Responses. Boston, Allyn and Bacon, 1992.

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

Johnson KL, Nicol T, Zecker SG, Kraus N: Developmental plasticity in the human auditory brainstem. J Neurosci 2008;28:4000–4007.

Johnson KL, Nicol TG, Zecker SG, Bradlow A, Skoe E, Kraus N: Brainstem encoding of voiced consonant-vowel stop syllables. Clin Neurophysiol, in press.

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

Klatt D: Software for cascade/parallel formant synthesizer. J Acoust Soc Am 1980;67:971–975.

Kraus N, Nicol T: Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends Neurosci 2005;28:176–181.

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

Levine RA, Liederman J, Riley P: The brainstem auditory evoked potential asymmetry is replicable and reliable. Neuropsychologia 1988;26:603–614.

Liu LF, Palmer AR, Wallace MN: Phase-locked responses to pure tones in the inferior colliculus. J Neurophysiol 2006;95:1926–1935.

Mathiak K, Hertrich I, Lutzenberger W, Ackermann H: Encoding of temporal speech features (formant transients) during binaural and dichotic stimulus application: a whole-head magnetencephalography study. Brain Res Cogn Brain Res 2000;10:125–131.

Morlet T, Collet L, Duclaux R, Lapillonne A, Salle B, Putet G, et al: Spontaneous and evoked otoacoustic emissions in pre-term and full-term neonates: is there a clinical application? Int J Pediatr Otorhinolaryngol 1995;33:207–211.

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

Oldfield RC: The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9:97–113.

Robinson K, Patterson RD: The duration required to identify the instrument, the octave, or the pitch chroma of a musical note. Music Percept 1995;13:1–15.

Russo N, Nicol T, Musacchia G, Kraus N: Brainstem response to speech syllables. Clin Neurophysiol 2004;115:2021–2030.

Russo NM, Bradlow AR, Skoe E, Trommer BL, Nicol T, Zecker S, Kraus N: Deficient brainstem encoding of pitch in children with autism spectrum disorders. Clin Neurophysiol 2008;119:1720–1731.

Schonwiesner M, Rubsamen R, von Cramon DY: Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. Eur J Neurosci 2005;22:1521–1528.

Schwartz J, Tallal P: Rate of acoustic change may underlie hemispheric specialization for speech perception. Science 1980;207:1380–1381.

Shtyrov Y, Kujala T, Lyytinen H, Kujala J, Ilmoniemi RJ, Naatanen R: Lateralization of speech processing in the brain as indicated by mismatch negativity and dichotic listening. Brain Cogn 2000;43:392–398.

Sidtis JJ: Predicting brain organization from dichotic listening performance: cortical and subcortical functional asymmetries contribute to perceptual asymmetries. Brain Lang 1982;17:287–300.

Sininger YS, Cone-Wesson B: Asymmetric cochlear processing mimics hemispheric specialization. Science 2004;305:1581.

Sininger YS, Cone-Wesson B: Lateral asymmetry in the ABR of neonates: evidence and mechanisms. Hear Res 2006;212:203–211.

Sininger YS, Cone-Wesson B, Abdala C: Gender distinctions and lateral asymmetry in the low-level auditory brainstem response of the human neonate. Hear Res 1998;126:58–66.

Smith JC, Marsh JT, Brown WS: Far-field recorded frequency following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 1975;39:465–472.

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

Spellacy F, Blumstein S: Ear preference for language and non-language sounds: a unilateral brain function. J Aud Res 1970;10:349–355.

Spivak LG, Seitz MR: Response asymmetry and binaural interaction in the auditory brain stem evoked response. Ear Hear 1988;9:57–64.

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

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

Zatorre RJ, Belin P: Spectral and temporal processing in human auditory cortex. Cereb Cortex 2001;11:946–953.

Zatorre RJB, Pascal P, Virginia B: Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002;6:37–46.