Clinical Neurophysiology 121 (2010) 1343–1350

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Rapid acoustic processing in the auditory brainstem is not related to cortical asymmetry for the syllable rate of speech

Daniel A. Abrams [a,*], Trent Nicol [a], Steven Zecker [a], Nina Kraus [a,b]

[a] *Auditory Neuroscience Laboratory, Department of Communication Sciences, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA*
[b] *Departments of Neurobiology and Physiology and Otolaryngology, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA*

ARTICLE INFO

Article history:
Accepted 25 February 2010
Available online 8 April 2010

Keywords:
Auditory cortex
Auditory brainstem
Children
Cerebral asymmetry
Speech

ABSTRACT

Objective: Temporal acuity in the auditory brainstem is correlated with left-dominant patterns of cortical asymmetry for processing rapid speech-sound stimuli. Here we investigate whether a similar relationship exists between brainstem processing of rapid speech components and cortical processing of syllable patterns in speech.
Methods: We measured brainstem and cortical evoked potentials in response to speech tokens in 23 children. We used established measures of auditory brainstem and cortical activity to examine functional relationships between these structures.
Results: We found no relationship between brainstem responses to fast acoustic elements of speech and right-dominant cortical processing of syllable patterns.
Conclusions: Brainstem processing of rapid elements in speech is not functionally related to rightward cortical asymmetry associated with the processing of syllable-rate features in speech. Viewed together with previous evidence linking brainstem timing with leftward cortical asymmetry for faster acoustic features, findings support the existence of distinct mechanisms for encoding rapid vs. slow elements of speech.
Significance: Results provide a fundamental advance in our knowledge of the segregation of sub-cortical input associated with cortical asymmetries for acoustic rate processing in the human auditory system. Implications of these findings for auditory perception, reading ability and development are discussed.

1. Introduction

The central auditory system consists of a highly complex network of sub-cortical and cortical nuclei characterized by an intricate pattern of connectivity between nuclei (Kaas and Hackett, 2000). Describing functional relationships between constituent areas of the auditory system brings us closer to understanding how humans process complex signals like speech and music.

Psychophysical research has demonstrated the importance of the temporal features of speech for speech understanding (Lisker and Abramson, 1964; Drullman et al., 1994; Shannon et al., 1995), and an important consideration is that the speech signal contains multiple temporal features that facilitate speech understanding (Rosen, 1992; Poeppel, 2003). Two categories of temporal features in speech that are critical for speech understanding are (1) the low-frequency temporal envelope, which we call the “speech envelope”, defined here as amplitude modulations in speech between 2 and 16 Hz (62.5–500 ms), and (2) more rapid modulations in the signal between 25 and 50 Hz (20–40 ms), which we will call the “segmental” temporal feature (Fig. 1). The speech envelope provides syllable pattern information, which is an essential cue for normal speech understanding (Drullman et al., 1994), while the segmental feature provides temporal information relevant for distinguishing between stop consonants (Lisker and Abramson, 1964).

Given the perceptual importance of the speech envelope and segmental temporal cues, there is considerable interest in understanding how the human auditory system responds to these temporal features. One line of research has used auditory evoked potentials in response to speech stimuli to describe the neural components that reflect the speech envelope and segmental features. With respect to the speech envelope, it was recently shown that this low-frequency temporal feature is processed predominantly in the right-hemisphere auditory cortex (Abrams et al., 2008, 2009). Specifically, it was shown that right-hemisphere cortical responses are of greater magnitude and follow the contours of

---

* Corresponding author. Address: Stanford Cognitive and Systems Neuroscience Laboratory, 780 Welch Road, Room 201, Palo Alto, CA 94304, USA. Tel.: +1 650 498 4206; fax: +1 650 736 7200.
E-mail address: daa@stanford.edu (D.A. Abrams).
URL: http://www.brainvolts.northwestern.edu (D.A. Abrams).

1388-2457/$36.00 © 2010 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2010.02.158

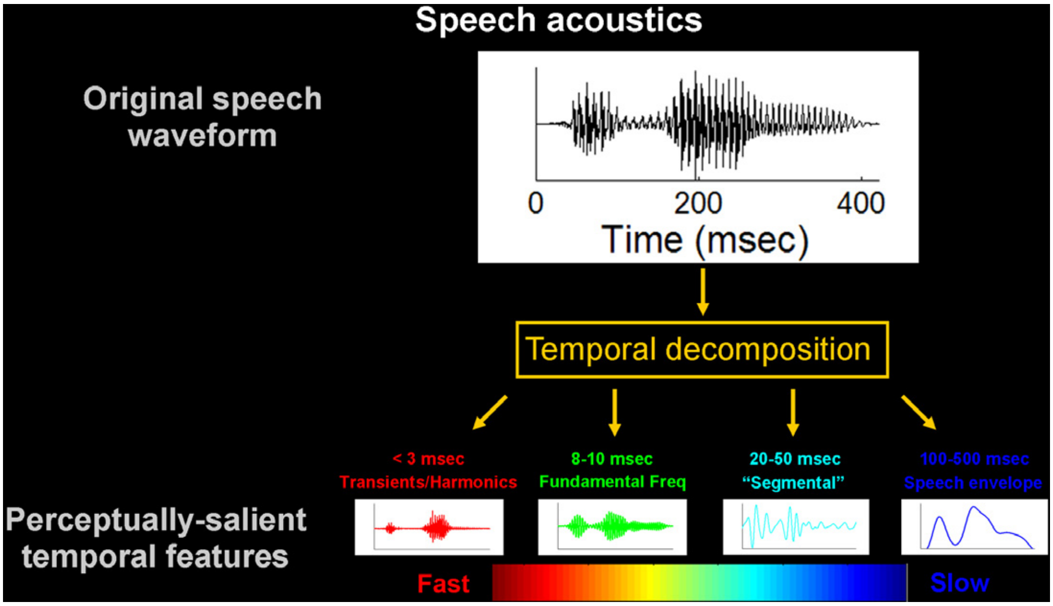

Fig. 1. Temporal decomposition of the speech signal into perceptually-salient temporal features natural speech can be decomposed based on temporal characteristics into each of these categories of speech features. The goal of the current study is to understand whether neural responses measured from the auditory brainstem and cortex that reflect these different temporal features are correlated with one another.

the speech envelope with greater precision relative to responses measured from left-hemisphere electrodes. With respect to the segmental temporal feature in speech, it has been shown in many studies that neural components measured in the human auditory brainstem closely mimic the temporal features of consonant–vowel stimuli (Galbraith et al., 1995; Akhoun et al., 2008; Johnson et al., 2008b; Hornickel et al., 2009).

A remaining question is whether neural components that reflect the speech envelope and segmental features are functionally related. This is an important question for a number of reasons. First, there is general interest in providing a more complete understanding of the functional relationships between the cortex, which appears to be important for speech envelope processing, and the brainstem, which appears to be important for processing the segmental feature of speech. Second, both the cortical processing of the speech envelope (Abrams et al., 2009) and brainstem processing of segmental features (Tzounopoulos and Kraus, 2009; Hornickel et al., 2009) have previously shown relationships with standardized behavioral measures of phonological processing and reading ability. Therefore, a more complete understanding of the relationship between these brain-based measures could provide insight into neural mechanisms underlying linguistic achievement.

To investigate this question, we measured auditory evoked potentials in a group of children in response to a variety of speech stimuli. To examine auditory brainstem function, we used a consonant–vowel stimulus that has been used in many studies (King et al., 2002; Wible et al., 2004; Skoe and Kraus, 2010; Kraus and Nicol, 2005; Banai et al., 2009; Dhar et al., 2009), and to measure cortical responses to the speech envelope, we used speech sentence stimuli that have elicited robust neural responses of this aspect of the speech signal in previous studies (Abrams et al., 2008, 2009). We then examined correlations between cortical measures of the speech envelope and brainstem measures of the segmental feature. We examined these measures in a group of children which included good, average, and poor readers as a means to provide a wide range of neurophysiologic profiles: poor readers have long been associated with abnormal cerebral asymmetry (Morgan, 1896), and more recently have demonstrated deficient responses to speech-sounds in the auditory brainstem (Cunningham et al., 2001; King et al., 2002; Wible et al., 2004; Banai et al., 2005, 2009; Hornickel et al., 2009; Chandrasekaran et al., 2009). An important consideration is that poor readers were included here to provide a neurophysiologically heterogeneous population, and were not included here to enable group comparisons between reading-impaired and typically developing children.

2. Methods

The research protocol was approved by the Institutional Review Board of Northwestern University. Parental consent and the child's assent were obtained for all evaluation procedures and children were paid for their participation in the study.

2.1. Participants

Participants consisted of 23 children between 9 and 15 years old who reported no history of neurological or otological disease and were of normal intelligence [scores >85 on either the Test of Nonverbal Intelligence (Brown et al., 1997) or Brief Cognitive Scale; (Woodcock and Johnson, 1977)].

Children were recruited from a database compiled in an ongoing project entitled Listening, Learning and the Brain. Children who had previously participated in this project and had indicated interest in participating in additional studies were contacted via telephone. Children were selected for this study to provide a cohort with a wide range of language abilities (Table 1) with the hope that a diversity of language abilities would provide a wide range of neurophysiologic profiles as shown in previous work (Abrams et al., 2006). Brainstem and cortical responses were measured during different sessions. Prior to all neurophysiological recordings, we performed a pure-tone hearing screening at 20 dB SPL for octaves between 500 and 4000 Hz on all subjects.

2.2. Cortical protocol

2.2.1. Stimuli

Stimuli consisted of the sentence stimulus “The young boy left home” produced in three modes of speech: conversational, clear

Table 1
Detailed subject information.

Subject	Age	Non-verbal IQ	Verbal IQ	Non-word reading	Single-word reading	Phonological processing
1	13.6	100	102	87	84	80
2	13.3	95	96	92	80	76
3	9.3	88	94	88	86	104
4	11.8	115	104	85	93	90
5	11.1	80	88	94	88	76
6	15.9	104	104	94	93	85
7	13.4	105	88	96	92	85
8	13.5	144	135	92	99	88
9	11.2	113	115	101	95	98
10	12.9	108	90	99	99	84
11	11.5	138	88	102	99	99
12	10.4	97	113	117	96	127
13	13.5	105	135	104	111	111
14	15.2	95	120	107	112	78
15	9.6	105	135	125	99	120
16	12.3	115	119	116	111	109
17	12.7	105	114	117	113	97
18	9.9	125	135	121	110	104
19	12.6	123	135	121	112	95
20	9.9	150	121	121	113	113
21	11.1	150	131	129	121	95
22	14	95	135	134	121	109
23	12.9	128	135	142	124	96

and compressed modes (Fig. 2, right). These stimuli have been described in previous works (Abrams et al., 2008, 2009). These three modes of speech have different speech envelope characteristics and were used as a means to elicit a variety of cortical activation patterns. Conversational speech is defined as speech produced in a natural and informal manner. Clear speech is a well-described mode of speech resulting from greater diction (Uchanski, 2005). Clear speech is naturally produced by speakers in noisy listening environments and enables greater speech intelligibility relative to conversational speech. There are many acoustic features that contribute to enhanced perception of clear speech relative to conversational speech, including greater intensity, slower speaking rate and more pauses. Most importantly with respect to the current work, an established feature of clear speech is greater temporal envelope modulations at lower modal frequencies than conversational speech, where modal frequency corresponds to the syllable rate (1–4 Hz) (Krause and Braida, 2004). Compressed speech approximates rapidly-produced speech and is characterized by a higher modal frequency. Compressed speech is more difficult to perceive compared to conversational speech (Beasley et al., 1980) and has been used in previous studies investigating cortical phase-locking to the speech envelope (Ahissar et al., 2001; Abrams et al., 2008, 2009).

Conversational and clear sentences were recorded at a sampling rate of 16 kHz in a soundproof booth by an adult male speaker. Conversational and clear speech sentences were equated for over-

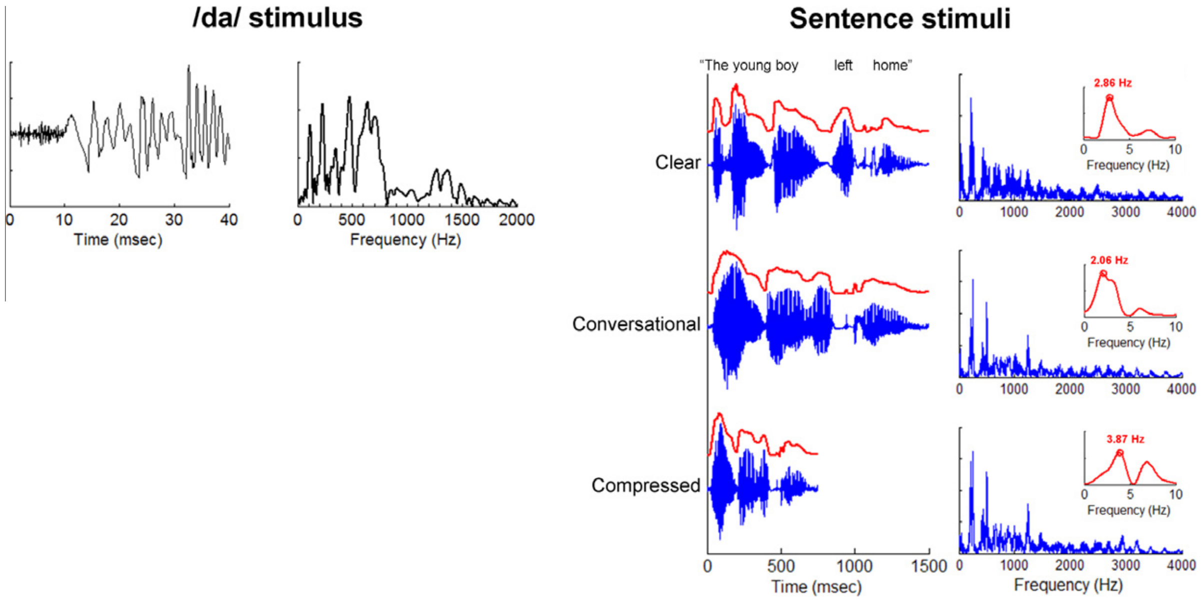

Fig. 2. Columns 1 and 2: brainstem stimulus "da" in the time and frequency domains. Column 3: cortical stimulus waveforms (blue) and broadband speech envelopes (red) for clear (top), conversational (center) and compressed (bottom) conditions. Greater amplitude envelope modulations are evident in the clear speech relative to the conversational stimuli. For example, there is no amplitude cue between "The" and "young" evident in the conversational stimulus envelope, however an amplitude cue is present in the clear stimulus envelope (0–450 ms post-stimulus onset). This phenomenon also occurs between the segments "boy" and "left" (450–900 ms post-stimulus onset). Column 4: frequency spectra for the three stimulus conditions (blue). Insets: frequency spectra of the speech envelopes (red) with the modal frequencies noted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

all duration to control for slower speaking rates in clear speech (Uchanski, 2005). This was achieved by compressing the clear sentence by 23% and expanding the conversational sentence by 23%. To generate the compressed sentence stimulus, we doubled the rate of the conversational sample using a signal-processing algorithm in Adobe Audition (Adobe Systems Inc.) which does not alter the pitch of the signal. The duration of the clear and conversational speech sentences was 1500 ms, and the duration of the compressed sentence was 750 ms.

2.2.2. Cortical recording and data processing procedures

All recording and data processing techniques used to describe cortical responses to the speech envelope are identical to those described in detail in two recent publications (Abrams et al., 2008, 2009). A PC-based stimulus delivery system (Neuroscan GenTask, Compumedics, Inc.) was used to output the sentence stimuli through a 16-bit converter at a sampling rate of 16 kHz. Speech stimuli were presented unilaterally to the right ear through insert earphones (Etymotic Research ER-2) at 80 dB SPL. Stimulus presentation was pseudorandomly interleaved. The polarity of each stimulus was reversed for half of the stimulus presentations to avoid stimulus artifacts in the cortical responses. Polarity reversal does not affect perception of speech samples (Sakaguchi et al., 2000). An interval of 1 s separated the presentation of sentence stimuli. Subjects were tested in a sound-treated booth and were instructed to ignore the sentences. To promote subject stillness during long recording sessions as well as diminish attention to the auditory stimuli, subjects watched a videotape movie of his or her choice and listened to the soundtrack to the movie in the non-test ear with the sound level set <40 dB SPL. This paradigm for measuring cortical evoked potentials has been used in previous studies investigating cortical asymmetry for speech-sounds (Bellis et al., 2000; Abrams et al., 2006) as well as other forms of cortical speech processing (Kraus et al., 1996; Banai et al., 2005; Wible et al., 2005). While it is acknowledged that cortical activity in response to a single stimulus presentation includes contributions from both the experimental speech stimulus and the movie soundtrack, auditory information in the movie soundtrack is highly variable throughout the recording session. Therefore, the averaging of auditory responses across 1000 stimulus presentations, which serves as an essential method for reducing the impact of noise on the desired evoked response, effectively removes contributions from the movie soundtrack. Cortical responses to speech stimuli were recorded with 31 tin electrodes affixed to an Electrocap (Electrocap International, Inc.) brand cap (impedance <5 KΩ). Additional electrodes were placed on the earlobes and superior and outer canthus of the left eye. These act as the reference and eye-blink monitor, respectively. Responses were collected (Neuroscan Acquire, Compumedics Inc.) at a sampling rate of 500 Hz for a total of 1000 repetitions each for clear, conversational and compressed sentences.

Processing of the cortical responses consisted of the following steps. First, excessively noisy segments of the continuous file (typically associated with subject movement) were manually rejected. The continuous file was high-pass filtered at 1 Hz and removal of eye-blink artifacts was accomplished using the spatial filtering algorithm provided by Neuroscan. Briefly, the spatial filtering algorithm performs a Principal Component Analysis decomposition of the blink artifact, and then creates a spatial filter that removes the artifact while retaining the EEG activity of interest. The continuous file was then low-pass filtered at 40 Hz to isolate cortical contributions and the auditory evoked potentials were then downsampled to a sampling rate of 200 Hz. All filtering was accomplished using zero phase-shift filters and downsampling was accompanied by IIR low-pass filtering to correct for aliasing (Compumedics, Inc.). The goal of this filtering scheme was to match the frequency range of the speech envelope (Rosen, 1992). Responses were epoched (2 s for responses to Clear and Conversational sentences; 1 s for responses to Compressed stimuli) and artifact rejected at a ±75 μV criterion. Responses were then subjected to noise reduction developed by our lab that has been used in improving the signal-to-noise ratio of brainstem and cortical evoked potentials (Abrams et al., 2008, 2009; Russo et al., 2009). The theoretical basis for the noise reduction is that it is assumed that auditory evoked potentials are largely invariant across individual stimulus repetitions while the background noise is subject to variance across stimulus repetitions. Thus, the mean evoked response is significantly diminished by the fraction of repetitions that least resembles it. If these noisy responses are removed, the signal-to-noise ratio of the cortical response improves considerably with virtually no change to morphology of the average waveform. For each subject, the algorithm first calculated the average time-domain response measured at a given electrode in response to a given stimulus condition (i.e., the mean amplitude waveform calculated across all repetitions for that stimulus). It then performed a Pearson's correlation between this average time-domain waveform and the waveform from each single repetition. The 30% of repetitions with the lowest Pearson's correlations from each stimulus condition were removed from subsequent analyses, and the remaining repetitions were averaged and re-referenced to a common reference computed across all electrodes. This process was repeated at all electrodes and all stimulus conditions. Therefore, following the artifact rejection and the noise reduction protocol, cortical responses from each subject represent the average of 650–700 repetitions of each stimulus.

2.2.3. Cortical data processing: measures of cortical speech envelope processing

All of the analysis techniques used to describe the cortical response to the speech envelope are identical to those described in detail in two recent publications (Abrams et al., 2008, 2009). All data analyses were performed using software written in Matlab (Version 7, The Mathworks, Inc.). Broadband stimulus envelopes were determined by performing a Hilbert transform on the broadband sentence waveforms (Drullman et al., 1994). The resulting amplitude envelopes were low-pass filtered at 40 Hz to isolate the speech envelope (Rosen, 1992) and resampled at 200 Hz to match sampling rate of the processed responses. We calculated the frequency of maximal power, known as the modal frequency (Ahissar et al., 2001), of the envelope of each speech sentence stimulus by performing a fast Fourier transforms of the low-pass filtered Hilbert envelope. FFTs were calculated using windows of 1 s and overlaps of 0.5 s, consistent with a previous report (Ahissar et al., 2001).

Data are presented for three temporal electrode pairs: (1) T3 and T4, (2) T5 and T6 and (3) Tp7 and Tp8 according to the modified International 10–20 recording system (Jasper, 1958). The modification is the addition of the Tp7–Tp8 electrode pair in which Tp7 is located midway between T3 and T5, and Tp8 is located midway between T4 and T6. Two types of analyses were performed on the data: cross-correlation and root mean squared (RMS) analyses, resulting in three neural measures in response to the speech envelope. First, cross-correlations between the broadband speech envelope and cortical responses at each temporal electrode for the "envelope-following period" (250–1500 ms for conversational and clear stimuli, 250–750 ms for the compressed stimulus) were performed using the "xcov" function in Matlab. The peak in the cross-correlation function, which represents the lag between stimulus and response, was found at each electrode between 50 and 150 ms, resulting in the first two measures. (1) *Phase-locking precision* was defined as the peak r-value and (2) *phase-locking timing* was defined as the lag at the peak r-value. R-values were Fisher-transformed prior to statistical analysis. Finally, (3) *RMS amplitudes* at each electrode were calculated for two different time ranges: the "onset" period was defined by the time ranges 0–250 ms for all stimuli; the "envelope-following" period was defined as 250–1500 ms for conversational and clear stimuli and 250–750 ms for the compressed stimulus.

2.2.4. Cortical data processing

A previous study showed that these measures of cortical envelope processing accounted for up to 44% of the variability in standardized reading scores and 50% in measures of phonological processing across a wide range of abilities (Abrams et al., 2009). Since the goal of the current work was to examine the relationship between auditory brainstem function and cortical asymmetry in processing of the speech envelope, we calculated these measures in individual subjects at specific temporal electrode sites to enable correlation analyses with auditory brainstem measures. Here are the measures of cortical speech envelope processing:

(1) Asymmetry of phase-locking *precision*: we first calculated mean left-hemisphere *r*-values (i.e., the mean of T3, T5 and Tp7) and right-hemisphere *r*-values (i.e., the mean of T4, T6 and Tp8) from the cross-correlation analysis of the compressed speech condition, and entered these values into the asymmetry index $(R - L)/(R + L)$.

(2) Asymmetry of phase-locking *timing*: we first calculated mean left-hemisphere lags (i.e., the mean of T3, T5 and Tp7) and right-hemisphere lags (i.e., the mean of T4, T6 and Tp8) from the cross-correlation analysis for each subject and each speech condition, and entered these values into the asymmetry index $(R - L)/(R + L)$.

(3) Asymmetry of phase-locking *magnitude*: we first calculated mean left-hemisphere amplitudes (i.e., the mean of T3 and Tp7) and right-hemisphere amplitudes (i.e., the mean of T4 and Tp8) from the analysis of the compressed speech condition, and then entered these values into the asymmetry index $(R - L)/(R + L)$. The reason T5 and T6 electrodes were not included in this measure is that a previous study found that responses measured at these electrodes did not show different patterns of asymmetry based on phase-locking magnitude (Abrams et al., 2009).

2.2.5. Brainstem protocol

The procedures to measure brainstem responses were identical to those that have been described (Russo et al., 2004; Skoe and Kraus, 2010). Brainstem responses were differentially recorded at a sampling rate of 20 kHz using a vertex electrode referenced to the right earlobe. The forehead served as ground. Three blocks of 1000 repetitions were collected at each polarity. Speech sounds were presented to the right ear at 80 dB sound pressure level (SPL) through insert earphones. The inter-stimulus interval was 51 ms. The stimulus used to evoke brainstem responses was the speech syllable /da/ (Fig. 2, left) synthesized at a sampling rate of 10 kHz. The stimulus was 40 ms in duration and consisted of five formants with an onset burst during the first 10 ms at F3, F4 and F5.

2.2.6. Data analysis

Analyses utilized previously described measures of the speech-evoked auditory brainstem response, including peak latency and amplitude measures, frequency domain and RMS measures, and measures describing parameters of the "VA onset complex" (Russo et al., 2004). Peaks V and A represent the onset of the speech-evoked brainstem response, and are thought to be analogous to peaks V and A in the traditional click-evoked ABR. The VA onset complex measures are derived from amplitude and latency values from peaks V and A, and include a series of four measures: inter-peak interval, amplitude, slope, and area (Russo et al., 2004; Skoe and Kraus, 2010). We also calculated composite brainstem scores derived from a subset of theoretically-important measures of the speech-evoked brainstem response. To calculate composite scores, each measure of brainstem function was first transformed into *Z*-scores and then averaged across measures. Five of the composite measures were recently described (Dhar et al., 2009) and consist of: (1) *brainstem onset response*, created from the latency of the two onset peaks, V and A; (2) *spectrotemporal features*, created from the latencies of peaks D, E and F; (3) *envelope boundary*, created from the latencies for peaks C and O, a period that corresponds to ~33 Hz (30 ms); (4) *pitch response*, calculated from the spectral magnitude of the brainstem response corresponding to the fundamental frequency (103–120 Hz) and the inter-peak intervals between peaks D and E, and E and F (8–10 ms); and (5) *harmonic response*, which consists of the spectral magnitude of the brainstem response corresponding to the first formant frequency (455–720 Hz) and the high-frequency components (721–1154 Hz) of the stimulus. Three additional measures published in slightly older studies (Russo et al., 2004; Abrams et al., 2006) were also included: *formant frequency response*, calculated as the mean spectral amplitude over the range of 220–1000 Hz (Russo et al., 2004); *acoustic transient response*, calculated from the latency of peaks V, A and O (Abrams et al., 2006), and *VA complex measures*, inter-peak interval, amplitude, and slope. The goal of using all of these measures is that they provide a thorough description of brainstem responses to temporal features in the transient/harmonic (<3 ms) and Fundamental Frequency (8–10 ms) ranges (Fig. 1). One of the measures, *envelope boundary*, provides temporal information within the "Segmental" range (20–50 Hz; 20–50 ms), as the timing difference between peaks C and O is ~30 ms.

Cortical measures consisted of the three previously described measures of speech envelope processing: asymmetry for phase-locking precision, phase-locking lag, and RMS amplitude asymmetry (Abrams et al., 2009). These particular measures are sensitive to standardized measures of literacy and phonological processing measured across individuals with a range of abilities.

We performed the following six analyses (in three categories) to examine potential relationships between auditory brainstem responses to speech-sounds and cortical responses to the speech envelope:

2.2.7. Analysis #1: correlation analyses

(1a) Pearson's correlations between the three measures of cortical speech envelope processing and individual measures of the speech-evoked auditory brainstem response (Russo et al., 2004; Skoe and Kraus, 2010). (1b) Pearson's correlations between the three measures of cortical speech envelope processing and the eight composite brainstem scores described above.

2.2.8. Analysis #2: cortical response comparisons, grouping based on brainstem response

(2a) Wilcoxon rank-sum test on cortical response measures for the top and bottom third ($n = 8$, each) of the cohort based on individual brainstem measures. (2b) Rank-sum test on cortical response measures for the top and bottom third based on composite brainstem scores.

2.2.9. Analysis #3: brainstem response comparisons, grouping based on cortical measures

(3a) Wilcoxon rank-sum test on individual brainstem measures for the top and bottom third based on cortical response measures. (3b) Rank-sum test on composite brainstem scores for the top and bottom third based on cortical response measures.

To prevent spurious results from correlation analyses, which can be overly biased by outlying data points, all raw brainstem and cortical values beyond 2 standard deviations (SD) of the mean were moved to the 2 SD point for that particular measure. Across cortical measures, 2 data points (out of a total of 69) were moved to the 2 SD point. For all analyses involving the comparison of means, we used the non-parametric Wilcoxon rank-sum test to examine group differences since there were only eight subjects in the two groups being compared (16 subjects total in this analysis). Pearson's correlations and rank-sum tests *p*-values <0.02 were considered statistically significant; this *p*-value cutoff represents an adjustment to account for the large number of comparisons being performed while not eliminating moderate brainstem-cortical correlations.

3. Results

3.1. Analysis #1: correlation analyses

Pearson's correlations were performed between individual measures of the speech-evoked auditory brainstem response (19 brainstem measures) and cortical responses to the speech envelope (3 measures). Results indicated one significant correlation (out of 57 correlations performed) between brainstem first formant frequency responses and cortical lag asymmetry. Furthermore, none of the Pearson's correlations between cortical responses to the speech envelope (3 measures) and eight composite measures of the speech-evoked auditory brainstem response met our criteria for significance.

3.2. Analysis #2: cortical response comparisons, grouping based on brainstem response

We performed Wilcoxon rank-sum tests on cortical response measures for the top and bottom third based on individual brainstem measures. Results from this analysis indicated no significant differences in any of the cortical measures between subjects in the top 33% and bottom 33% for a given brainstem measure. When we used the eight composite brainstem scores to define the groups rather than individual measures, again there were no significant differences in any of the cortical measures between subjects in the top 33% and bottom 33%.

3.3. Analysis #3: brainstem response comparisons, grouping based on cortical measures

We performed Wilcoxon rank-sum tests on individual brainstem measures for the top and bottom third based on cortical response measures. Results indicated one significant rank-sum result (out of 57 rank-sum tests in which grouping was based on cortical measures) between brainstem frequency response to F1 over the FFR period and cortical Lag asymmetry ($p = 0.015$). There were no other significant differences when we performed Wilcoxon rank-sum tests on composite brainstem measures for the top and bottom third based on cortical response measures.

4. Discussion

We investigated whether there is a relationship between rapid acoustic processing in the human auditory brainstem and cortical processing of the slowly-varying temporal envelope of speech. We failed to see any systematic relationship between the auditory brainstem response and cortical processing of the speech envelope. Analyses included Pearson's correlations between established measures of brainstem and cortical responses as well as the grouping of subjects based on brainstem and cortical responses; in all but two cases, results failed to reach statistical significance, or even show trends for significance. Based on the consistency of these negative results, results strongly suggest that there is no relationship between the fidelity of auditory processing of transients and fundamental frequency in the auditory brainstem and cortical asymmetry for syllable-rate processing of speech.

One brainstem-cortical relationship showed statistical significance: the relationship between brainstem response to F1 frequency component and cortical lag asymmetry. We argue that this significant correlation does not represent a convincing scientific finding given the number of comparisons and correlations performed in this study. Furthermore, this particular relationship is not easily explained within the existing framework describing brainstem-cortical interactions since the brainstem response to F1 had not previously been implicated as a measure that is sensitive to cortical function (Banai et al., 2005; Wible et al., 2005; Abrams et al., 2006).

4.1. Specificity for temporal acuity in the auditory brainstem and speech-related cerebral asymmetry

The current experiment was designed and performed in an extremely similar manner as a previous work examining the relationship between brainstem responses and cortical asymmetry for rapid features in speech (Abrams et al., 2006). Similarities include similar mixed populations of normal and reading-impaired children, as well as identical brainstem methods with respect to stimulus, collection and processing techniques. Furthermore, both studies utilized similar cortical methods, with the only substantial difference being the stimuli, which, in the former study, were designed to maximize rapid acoustic processing in the left-hemisphere (the syllable /da/) and in the present study, to maximize slow acoustic processing in the right-hemisphere (complete sentences). The brainstem-cortical analyses were also extremely similar between these two studies in that they quantified the relationship between brainstem measures and study-specific measures of cortical asymmetry, including asymmetry of cortical RMS amplitude.

Despite the consistent study design, data collection and analyses, the results from these two studies are in stark contrast: results indicate that auditory brainstem timing for rapid acoustic features is only related to cortical asymmetry that reflects rapid components of the speech signal and not to slow features. Why might this be? One possibility is that rapid acoustic features in speech may represent a greater processing load for the auditory system compared with slower features, and to accommodate this increased load, the processing of rapid acoustic features in speech may require greater neural synchrony between the constituent nuclei in the auditory system relative to slower acoustic features.

Another possibility is that the auditory brainstem is actually related to right-dominant speech envelope responses in cortex, but the specific measures of brainstem function described here are not sensitive to this aspect of cortical function. For example, we filter out low frequencies below 100 Hz from evoked potentials as a means of isolating brainstem responses, and lower frequency elements conceivably generated by the brainstem could relate to cortical processing of the syllable rate of speech. If brainstem function is actually correlated with rightward asymmetry for the speech envelope in a manner that we were not able to describe in the current work, this would suggest exquisite response specificity for speech envelope processing. This is based on the fact that, over the last decade, the Kraus lab has thoroughly characterized the auditory brainstem response to the speech-sound /da/ in both time and frequency domains (Cunningham et al., 2001; King et al., 2002; Russo et al., 2004, 2005; Wible et al., 2004; Skoe and Kraus 2010; Chandrasekaran et al., 2009; Hornickel et al., 2009; Dhar et al., 2009), and in the current work none of these measures showed a reliable relationship to cortical speech envelope asymmetry. Therefore, if a currently-undiscovered measure of auditory brainstem processing is actually related to cortical envelope asymmetry, it is likely a highly specific relationship involving a brainstem measure that is uncorrelated with other measures of brainstem function. It is hoped that future studies can further investigate this question, perhaps using a longer speech stimulus compared to the one used here which would enable a more robust metric of steady-state response processing in the auditory brainstem.

4.2. Brainstem mechanisms for temporal processing of speech

Results described here add to a growing body of literature describing brainstem processing of acoustically-complex signals such as speech and music. While historically the auditory brainstem was thought to play a relatively passive role in the processing of biologically-important signals, recent studies have demonstrated that human brainstem activity reflects aspects of perception (Johnson et al., 2007; Song et al., 2008b; Hornickel et al., 2009; Parbery-Clark et al., 2009), development (Johnson et al., 2008a), musical experience (Musacchia et al., 2007; Wong et al., 2007; Kraus et al., 2009; Lee et al., 2009; Strait et al., 2009; Tzounopoulos and Kraus, 2009; Parbery-Clark et al. 2009; Bidelman et al., 2009; Chandrasekaran and Kraus, 2010), language experience (Krishnan et al., 2005, 2009a,b; Russo et al., 2005) and critical language skills (King et al., 2002; Wible et al., 2004; Banai et al., 2005, 2009; Song et al., 2008a; Hornickel et al., 2009; Chandrasekaran et al., 2009). The current study adds to this literature by showing that, despite the well-documented sensitivity of the auditory brainstem response to a host of perceptual, cognitive and experiential factors, sub-syllabic spectrotemporal processing in the brainstem does not appear to be related to cortical specialization for the speech envelope. This is a surprising result for two reasons. First, the speech envelope is essential for the normal perception of speech (Drullman et al., 1994; Shannon et al., 1995), and while the current study did not examine perception, a previous study showed a strong link between speech perception and cortical responses similar to those described here (Ahissar et al., 2001). Second, the speech envelope, which encompasses the syllable rate of speech (Rosen, 1992), is thought to play an important role in development: an influential theory of reading acquisition argues that sensitivity to syllables in young children precedes sensitivity to phonemes, and therefore the mastering of syllable-level skills is paramount to phonological development (Stanovich, 1992; Anthony et al., 2003; Ziegler and Goswami, 2005). Taken together, the speech envelope is thought to play a critical role in both perception and development, yet the auditory brainstem – at least for the brainstem measures assessed here – appears to be functionally unrelated to cortical speech envelope processing. The current results highlight the specificity of the auditory brainstem processing to other aspects of perception, cognition and experience. A critical consideration is that only certain features of the speech-evoked brainstem response are related to behavioral impairments and expertise (Kraus and Nicol, 2005; Krizman et al., 2010).

4.3. Cortical mechanisms for temporal processing of speech

Findings described in the current paper, in conjunction with a previous work investigating brainstem timing and cortical rate asymmetry (Abrams et al., 2006), enable speculation on the mechanisms hypothesized to serve acoustic rate asymmetries in auditory cortex (Poeppel, 2003). Very little is known about the nature of this mechanism. One possibility is that a single neural mechanism is responsible for routing rapid and slow temporal processing to the left and right auditory cortices, respectively. If this were the case, one might hypothesize that auditory brainstem function would be related to both rapid and slow forms of cortical asymmetry in a similar manner. Alternatively, it is plausible that there are two separate cortical mechanisms, a rapid processing mechanism and a slow processing mechanism, and that these separable mechanisms are characterized by different network properties. The accumulated findings suggest the latter scenario since brainstem-cortical relationships appear to differ between the processing of rapid and slow temporal features in speech. It is hoped that future studies in human and animal models can more directly address the mechanisms serving acoustic rate asymmetries described here.

4.4. Implications for reading and phonological processing

Accumulated results suggest that neural synchrony among constituent areas of the auditory system could serve as a mechanism necessary for normal speech perception, reading acquisition, and phonological abilities (Banai et al., 2005, 2009; Basu et al., 2010; Wible et al., 2005; Abrams et al., 2006; Chandrasekaran et al., 2009; Dhar et al., 2009; Hornickel et al., 2009), however results from the current study failed to show synchrony between consonant–vowel processing in the brainstem and cortical processing of the speech envelope. One interpretation of this finding is that impaired synchrony described in other works is a somewhat specific phenomenon as it relates to perceptual and reading-related abilities. For example, a trend in these previous studies is that abnormal perception and reading-related skills are associated with impaired brainstem-cortical synchrony for processing of rapid acoustical features (e.g., processing of speech transients and brief consonant–vowel stimuli). Perhaps the exquisite neural timing necessary for processing and synchronizing to these rapid acoustical features imposes a specific limitation on the reading-impaired auditory system. An exciting future study could address whether dyslexic children with poor brainstem timing and atypical leftward asymmetry to fast speech cues constitute the same population who demonstrate atypical rightward asymmetry to speech envelope cues. The results of this study could be revealing: if these subjects represent the same population, it would demonstrate the existence of a pervasive, yet mechanistically distinct, auditory timing disorder; if these subjects represent separate populations, results could provide an objective way to distinguish distinct underlying deficits contributing to reading disorders.

A caveat for this work is that one cannot definitively conclude that there is no relationship between brainstem and cortical envelope responses from a set of null results. Nevertheless, we believe that a number of methodological and analytical considerations enable a strong suggestion of this possibility. First, brainstem and cortical evoked potentials are thought to be extremely robust and stable within subjects, and effects are generally apparent in small groups (Cunningham et al., 2001; Wible et al., 2002, 2004; Krishnan et al., 2005; Wible et al., 2005; Abrams et al., 2008, 2009; Bidelman et al., 2009; Bidelman and Krishnan, 2009; Hornickel et al., 2009; Skoe and Kraus, 2010). Second, our data analysis examined relationships between 19 different measures of brainstem function and three cortical measures, and included a variety of statistical techniques to probe potential relationships (correlations, *t*-tests, and Wilcoxon rank-sum tests). Results showed that only two brainstem-cortical relationships reached our moderate criteria for significance ($p < 0.02$), which is fewer than the number of statistically significant relationships predicted by chance ($\sim$3). Finally, with respect to the number of subject participants in this work ($n = 23$), a power analysis shows that we had the ability to detect effect sizes of $r = 0.5$ with a power of $\sim$65% ($\alpha = 0.02$). While we recognize that this is not an ideal power value, we do not believe that it is prohibitively low.

In conclusion, we have shown that there is no demonstrable relationship between auditory brainstem responses to syllabic elements and cortical asymmetry for slow features in speech. These results complement previous findings that show a significant relationship between auditory brainstem timing and cortical asymmetry for rapid acoustic features in speech. Results suggest a dichotomy between neural mechanisms serving acoustic rate asymmetries in auditory cortex.

Competing financial interest

The authors declare no competing financial interests.

Acknowledgements

We thank the children who participated in this study and their families and members of the Auditory Neuroscience Lab for data collection and processing. This work is supported by the National Institutes of Health Grant R01 DC01510-10 and National Organization for Hearing Research Grant 340-B208.

References

Abrams DA, Nicol T, Zecker SG, Kraus N. Auditory brainstem timing predicts cerebral asymmetry for speech. J Neurosci 2006;26:11131–7.

Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. J Neurosci 2008;28:3958–65.

Abrams DA, Nicol T, Zecker S, Kraus N. Abnormal cortical processing of the syllable rate of speech in poor readers. J Neurosci 2009;29:7686–93.

Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. Proc Natl Acad Sci USA 2001;98:13367–72.

Akhoun I, Gallego S, Moulin A, Menard M, Veuillet E, Berger-Vachon C, et al. The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. Clin Neurophysiol 2008;119:922–33.

Anthony JL, Lonigan CJ, Driscoll K, Phillips BM, Burgess SR. Phonological sensitivity: a quasi-parallel progression of word structure units and cognitive operations. Read Res Quart 2003;38:470–87.

Banai K, Nicol T, Zecker SG, Kraus N. Brainstem timing: implications for cortical processing and literacy. J Neurosci 2005;25:9850–7.

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. Cereb Cortex 2009.

Basu M, Krishnan A, Weber-Fox C. Brainstem correlates of temporal auditory processing in children with specific language impairment. Dev Sci 2010;13:77–91.

Beasley DS, Bratt GW, Rintelmann WF. Intelligibility of time-compressed sentential stimuli. J Speech Hear Res 1980;23:722–31.

Bellis TJ, Nicol T, Kraus N. Aging affects hemispheric asymmetry in the neural representation of speech sounds. J Neurosci 2000;20:791–7.

Bidelman GM, Krishnan A. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. J Neurosci 2009;29:13165–71.

Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci 2009.

Brown L, Sherbenou R, Johnsen S, editors. Test of nonverbal intelligence: a language-free measure of cognitive ability. Austin, TX: Pro-Ed; 1997.

Chandrasekaran B, Kraus N. Music, noise-exclusion, and learning disabilities. Music Percept 2010;27:297–306.

Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N. Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. Neuron 2009;64:311–9.

Cunningham J, Nicol T, Zecker S, Bradlow A, Kraus N. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol 2001;112:758–67.

Dhar S, Abel R, Hornickel J, Nicol T, Skoe E, Zhao W, et al. Exploring the relationship between physiological measures of cochlear and brainstem function. Clin Neurophysiol 2009;120:959–66.

Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. J Acoust Soc Am 1994;95:1053–64.

Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM. Intelligible speech encoded in the human brain stem frequency-following response. Neuroreport 1995;6:2363–7.

Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. Proc Natl Acad Sci USA 2009;106:13022–7.

Jasper HH. The ten–twenty electrode system of the international federation. Electroencephalogr Clin Neurophysiol 1958;10:371–5.

Johnson KL, Nicol TG, Zecker SG, Kraus N. Auditory brainstem correlates of perceptual timing deficits. J Cogn Neurosci 2007;19:376–85.

Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. J Neurosci 2008a;28:4000–7.

Johnson KL, Nicol T, Zecker SG, Bradlow AR, Skoe E, Kraus N. Brainstem encoding of voiced consonant–vowel stop syllables. Clin Neurophysiol 2008b;119:2623–35.

Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA 2000;97:11793–9.

King C, Warrier CM, Hayes E, Kraus N. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett 2002;319:111–5.

Kraus N, Nicol T. Brainstem origins for cortical ‘what’ and ‘where’ pathways in the auditory system. Trends Neurosci 2005;28:176–81.

Kraus N, McGee TJ, Carrell TD, Zecker SG, Nicol TG, Koch DB. Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science 1996;273:971–3.

Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. Ann NY Acad Sci 2009;1169:543–57.

Krause JC, Braida LD. Acoustic properties of naturally produced clear speech at normal speaking rates. J Acoust Soc Am 2004;115:362–78.

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;25:161–8.

Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. J Cogn Neurosci 2009a;21:1092–105.

Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. Experience-dependent neural representation of dynamic pitch in the brainstem. Neuroreport 2009b;20:408–13.

Krizman J, Skoe E, Kraus N. Stimulus rate and subcortical auditory processing of speech. Audiol Neurootol 2010;15:332–42.

Lee KM, Skoe E, Kraus N, Ashley R. Selective subcortical enhancement of musical intervals in musicians. J Neurosci 2009;29:5832–40.

Lisker L, Abramson AS. A cross-langauge study of voicing in initial stops: acoustical measurements. Word 1964;20:384–422.

Morgan WP. A case of congenital word-blindness. Brit Med J 1896;2:1378.

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;104:15894–8.

Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci 2009;29:14100–7.

Poeppel D. The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. Speech Commun 2003;41:245–55.

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond B Biol Sci 1992;336:367–73.

Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clin Neurophysiol 2004;115:2021–30.

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

Russo N, Zecker S, Trommer B, Chen J, Kraus N. Effects of background noise on cortical encoding of speech in autism spectrum disorders. J Autism Dev Disord 2009;39:1185–96.

Sakaguchi S, Arai T, Murahara Y. The effect of polarity inversion of speech on human perception and data hiding as an application. International conference on acoustics, speech, and signal processing 2000;2:917–20.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science 1995;270:303–4.

Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear Hear 2010. doi:10.1097/AUD.0b013e3181cdb272.

Song JH, Banai K, Kraus N. Brainstem timing deficits in children with learning impairment may result from corticofugal origins. Audiol Neurootol 2008a;13:335–44.

Song JH, Skoe E, Wong PC, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 2008b;20:1892–902.

Stanovich KE. Speculations on the causes and consequences of individual differences in early acquisition. In: Gough PB, Ehri LE, Treiman R, editors. Reading acquisition. Hillsdale, NJ: Erlbaum; 1992. p. 307–42.

Strait DL, Kraus N, Skoe E, Ashley R. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. Eur J Neurosci 2009;29:661–8.

Tzounopoulos T, Kraus N. Learning to encode timing: mechanisms of plasticity in the auditory brainstem. Neuron 2009;62:463–9.

Uchanski RM. Clear Speech. In: Pisoni DB, Remez RE, editors. Handbook of Speech Perception. Malden, MA: Blackwell Publishers; 2005.

Wible B, Nicol T, Kraus N. Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. Clin Neurophysiol 2002;113:485–94.

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.

Wible B, Nicol T, Kraus N. Correlation between brainstem and cortical auditory processes in normal and language-impaired children. Brain 2005;128:417–23.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 2007;10:420–2.

Woodcock R, Johnson M, editors. Woodcock-Johnson psychoeducational battery: tests of cognitive ability. Allen, TX: DLM Teaching Resources; 1977.

Ziegler JC, Goswami U. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. Psychol Bull 2005;131:3–29.