Subcortical representation of speech fine structure relates to reading ability
Jane Hornickel\textsuperscript{a}, Samira Anderson\textsuperscript{a}, Erika Skoe\textsuperscript{a}, Han-Gyol Yi\textsuperscript{a} and Nina Kraus\textsuperscript{a,b}

Impaired perception of consonants by poor readers is reflected in poor subcortical encoding of speech timing and harmonics. We assessed auditory brainstem representation of higher harmonics within a consonant-vowel formant transition to identify relationships between speech fine structure and reading. Responses were analyzed in three ways: a single stimulus polarity, adding responses to inverted polarities (emphasizing low harmonics), and subtracting responses to inverted polarities (emphasizing high harmonics). Poor readers had a reduced representation of higher speech harmonics for subtracted polarities and a single polarity. No group differences were found for the fundamental frequency. These findings strengthen the evidence of subcortical encoding deficits in poor readers for speech fine structure and delineate effective strategies for capturing these neural impairments in humans. NeuroReport 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, Illinois, USA
Department of Communication Sciences, Auditory Neuroscience Laboratory and Departments of Neurobiology and Physiology, Northwestern University Otolaryngology, Evanston, Illinois, USA

Tel: +1 847 491 2459; fax: +1 847 491 4975; e-mail: jhornickel@northwestern.edu

Keywords: auditory brainstem, electrophysiology, evoked potentials, reading, speech

Introduction
Children with reading impairments often perform poorly on psychophysical auditory processing tasks relative to typically developing peers [1] and auditory processing skills in infancy or preschool can predict later language and reading ability [2,3]. Deficits in auditory processing likely contribute to the well-documented impairments in discriminating and using stop consonants by children with reading disorders [4-6] due to the transient elements and rapidly changing frequency sweeps (i.e., formants) characterizing stop consonants [7]. Poor perception and misuse of stop consonants by children with reading impairments may be due to a combination of an impaired neural representation of stimulus features and an inability to use regularly occurring acoustic cues in their environment to modulate sensory representation (i.e., the particular temporal and spectral features that represent linguistic information [8-10]).

Consistent with behavioral deficits in auditory processing, many children with reading disorders exhibit impaired subcortical encoding of speech and speech-like signals relative to good readers [11-16]. These impairments are particularly evident for the acoustic cues important for distinguishing stop consonants, namely, the neural representation of stimulus timing and higher harmonics (including those that fall in the range of speech formants). However, the neural representation of the fundamental frequency (an important pitch cue) is not disrupted in poor readers. Children with poor reading ability or language skills have significantly slower neural response timing, less robust neural encoding of formant-related stimulus harmonics, and less robust tracking of frequency contours compared with typically developing children, and the neural responses to these acoustic cues are predictive of reading across children with a wide range of ability [11-16]. Although the relationship between neural response timing and reading has been established with multiple speech and speech-like stimuli, further exploration is required to understand the nature of the deficits seen for poor readers in the encoding of stimulus harmonics. Speech harmonics are critical cues for perceiving stop consonants [7] and their neural representation is likely to be tightly related to reading ability.

Deficits in the representation of formant-related harmonic encoding have been found using an abbreviated speech stimulus (formant transition only), for which responses to inverted stimulus polarities (180° out of phase) were added [11,12]. This adding of polarities accentuates the lower frequency components in the response that contribute to the amplitude envelope, including the fundamental frequency (F\(_0\)) [17-19]. In the current analysis, we used a second processing method of subtracting the responses to inverted polarities to extract the components of the response reflecting the stimulus fine structure (i.e., formant-related harmonics) [17,20]. Unlike previous work with an isolated formant transition, here we measured responses to a full consonant-vowel syllable, which was longer, contained a full vowel, and was consequently presented at a slower rate than the isolated formant transition stimulus. Owing to the slower presentation rate and full vowel of the consonant-vowel syllable, we chose to additionally utilize...
the method of subtracting polarities because we anticipated it would be more fruitful in revealing group differences through the maintenance of stimulus fine structure in the response. In addition, we analyzed responses to a single stimulus polarity, in which both the amplitude envelope and the fine structure would be represented. We compared poor readers with average readers in their encoding of the F₀ and harmonics using the two processing techniques and responses to an individual polarity. We hypothesized that deficits in the neural representation of formant-related harmonics account in part for reading difficulties in children. We predicted that these deficits in processing speech harmonics would be more evident in analysis techniques that emphasize or maintain the fine structure response elements (subtracting polarities and a single polarity) rather than those which emphasize envelope-related elements (adding polarities).

**Methods**

Fifty-one children aged 8–13 years (mean = 10.7, 16 girls) participated in this study. All had normal hearing (air-conduction and bone-conduction thresholds < 20 dB hearing level for octaves from 250–8000 Hz), click-evoked auditory brainstem responses within normal limits (100 µs presented at 31.3 Hz), and full-scale IQ scores above 85 on the Wechsler Abbreviated Scale of Intelligence [21]. Reading ability was assessed using the Test of Silent Word Reading Fluency [22]. Poor readers were defined as children with scores of less than or equal to 90 (n = 25; eight girls). All other participants had Test of Silent Word Reading Fluency scores within the average range (95–105; n = 26; eight girls). All procedures were approved by the Northwestern University Institutional Review Board and the children and a parent/guardian gave their informed assent and consent, respectively.

**Stimuli and recording parameters**

All stimuli, presentation parameters and recording techniques have been used previously [11,14]. The speech stimulus was a synthesized 170-ms [da] syllable, with dynamic first, second, and third formants during the first 50 ms, presented monaurally using inverted polarity (95–105; n = 26; eight girls). All other participants had Test of Silent Word Reading Fluency scores within the average range (95–105; n = 25; eight girls). All procedures were approved by the Northwestern University Institutional Review Board and the children and a parent/guardian gave their informed assent and consent, respectively.

**Add and subtract processing techniques**

Polarity-specific averages were analyzed individually (single polarity), added (add), or subtracted (subtract). The add technique minimizes phase-dependent activity, including artifact from the transducer and cochlea, whereas the phase-independent component of the neural response remains. The add technique also particularly emphasizes the response to lower frequency harmonics, while minimizing the response to the high-frequency stimulus fine structure [17–19]. In contrast, the subtract technique retains the response to the fine structure of the signal and minimizes the envelope response by correcting the half-wave rectification occurring in cochlear transduction [17,20]. Although the subtract technique does not reduce cochlear or transducer artifact, our testing protocol minimizes artifact contamination by using transducers with tube earphones, common mode rejection referencing, and alternating polarities [23], and we have not observed artifact in this dataset or others. The single-polarity response reflects both the amplitude envelope and the fine structure of the stimulus (see Fig. 1a–d).

**Data analyses**

Fast-fourier transforms were conducted over the response to the stimulus formant transition (20–60 ms). Spectral amplitudes were averaged over 50–150 Hz (F₀), 450–750 Hz (mid harmonics), and 750–1150 Hz (high harmonics). We selected these regions because group differences between good and poor readers using a 40-ms formant transition occur for 410–755 and 750–1150 Hz, but not for lower harmonic ranges that are likely important for the perception of pitch [11,24].

The poor readers were compared with the average readers on the spectral magnitudes of their responses in the three frequency ranges (F₀, low harmonics, high harmonics) for each of the three analysis techniques (single polarity, add, subtract) using Independent t-tests. Multiple comparisons were corrected for by setting α to be 0.005.

**Results**

Poor readers had significantly smaller response amplitudes for the high harmonics when using the subtract technique (t₄₉ = 3.07, P = 0.004) and this effect was trending for the single-polarity response (t₄₉ = 2.56, P = 0.014). These group differences were not revealed using the add technique (t₄₉ = 1.33, P = 0.19), which suggests that these higher harmonic effects are magnified by the subtract technique and single-polarity responses, which more robustly represent stimulus fine structure (see Fig. 1f–g). No group differences were found for the F₀ (single polarity: t₄₉ = 0.62, P = 0.54; add: t₄₉ = 0.54, P = 0.60; subtract: t₄₉ = –1.49, P = 0.14) or for the mid harmonics (single polarity: t₄₉ = 1.49, P = 0.14; add: t₄₉ = 0.35, P = 0.73; subtract: t₄₉ = 1.15, P = 0.26) in any condition.

**Discussion**

The current results reinforce and expand upon previous findings that brainstem encoding of formant-related harmonics is impaired in poor readers. Poor readers had weaker representation of the high harmonics than average readers, revealed using the subtract technique and when
analyzing responses to a single polarity. Group differences were not found for the neural representation of the $F_0$ for any analysis technique, highlighting that neural encoding deficits in poor readers are restricted to speech fine structure important for distinguishing consonants.

We saw no group differences in the representation of speech fine structure for the add method likely due to an interaction between the stimulus characteristics and processing techniques. As the formant transition of the full consonant-vowel syllable is longer than the previously utilized formant transition stimulus, changes in formant frequencies are slower. Although the duration and presentation rate of this syllable-length stimulus was rapid enough to be difficult for children with reading and language impairments to perceive [25], the slightly longer formant transition and the slower presentation rate used here may have not taxed the auditory system to the same extent as the formant transition stimulus, minimizing group differences in response to the speech envelope. As the add technique highlights the amplitude envelope and low harmonic component of the response and minimizes the contribution of the fine structure elements [17,20], it is not surprising that group differences in harmonic encoding were not found with the add processing technique. However, when the same responses are analyzed using the subtract technique, group differences in encoding of formant-related high harmonics did emerge. The subtract technique emphasizes the components of the response relating to the stimulus fine structure, which directly reflects the speech formants [17,19,20]. Group differences were also trending for responses to a single polarity, which contains response elements reflecting both the stimulus amplitude enve-
ope and fine structure. Thus, the subtract technique most effectively revealed group differences in high-harmonic representation by minimizing the representation of the amplitude envelope and emphasizing the representation of the stimulus fine structure in response to a full consonant-vowel syllable.

Conclusion
Overall, poor readers had weaker encoding of stimulus fine structure than average readers, with no group differences for the F0, replicating previous results [11,12]. The present results lend further support for the hypothesis that reading ability is linked to auditory processing. Recent analytical modeling has shown that brainstem measures representing timing and harmonic elements significantly predict variance in reading ability, even when taking into account phonological awareness [16]. Impairments in the neural encoding of acoustic elements crucial for differentiating consonants, such as formant frequencies (reflected in speech fine structure), may contribute to poor phonological development and the poor consonant differentiation and phonological skills seen in children with reading impairments. Overall, these results highlight the contribution of auditory processing to learning and communication skills and reveal the effectiveness of different strategies for capturing these neural signatures in humans.

Acknowledgements
The authors thank Jennifer Krizman, Alexandra Parbery-Clark, and Trent Nicol for their review of the manuscript, the members of the Auditory Neuroscience lab for their assistance with data collection, and the children and their families for participating.

This work was supported by National Institutes of Health (R01DC015110) and the Hugh Knowles Center at Northwestern University.

Conflicts of interest
There are no conflicts of interest.

References