

Neural Timing Is Linked to Speech Perception in Noise

Samira Anderson,^{1,3} Erika Skoe,^{1,3} Bharath Chandrasekaran,^{1,2} and Nina Kraus^{1,3,4,5}

¹Auditory Neuroscience Laboratory, ²Communication Neural Systems Group, Departments of ³Communication Sciences, ⁴Neurobiology and Physiology, and ⁵Otolaryngology, Northwestern University, Evanston, Illinois 60208

Understanding speech in background noise is challenging for every listener, including those with normal peripheral hearing. This difficulty is attributable in part to the disruptive effects of noise on neural synchrony, resulting in degraded representation of speech at cortical and subcortical levels as reflected by electrophysiological responses. These problems are especially pronounced in clinical populations such as children with learning impairments. Given the established effects of noise on evoked responses, we hypothesized that listening-in-noise problems are associated with degraded processing of timing information at the brainstem level. Participants (66 children; ages, 8–14 years; 22 females) were divided into groups based on their performance on clinical measures of speech-in-noise (SIN) perception and reading. We compared brainstem responses to speech syllables between top and bottom SIN and reading groups in the presence and absence of competing multitalker babble. In the quiet condition, neural response timing was equivalent between groups. In noise, however, the bottom groups exhibited greater neural delays relative to the top groups. Group-specific timing delays occurred exclusively in response to the noise-vulnerable formant transition, not to the more perceptually robust, steady-state portion of the stimulus. These results demonstrate that neural timing is disrupted by background noise and that greater disruptions are associated with the inability to perceive speech in challenging listening conditions.

Introduction

Speech consists of rapidly changing elements that require fine-grained neural representation of temporal information, especially in background noise. Temporal cues are important components of auditory object formation (Shinn-Cunningham and Best, 2008), a necessary element of auditory stream segregation. Early stages of auditory stream segregation occur subcortically (Pressnitzer et al., 2008; Parbery-Clark et al., 2009a), and timing cues needed for speech perception and auditory stream segregation are preserved in the brainstem via neural synchrony (Kraus and Nicol, 2005; Akhoun et al., 2008; Hornickel et al., 2009; Tzounopoulos and Kraus, 2009). It is well established that neural synchrony is degraded in noise, leading to delayed and reduced auditory evoked responses from cortical (Warrier et al., 2004; Billings et al., 2009; Russo et al., 2009) and brainstem structures (Hall, 1992; Cunningham et al., 2001; Burkard and Sims, 2002; Russo et al., 2004). In the auditory brainstem response (ABR), background noise disrupts the representation of temporal aspects of the stimulus, leading to delayed scalp-recorded far field responses to the time-varying features of a speech stimulus (e.g., the onset and formant transition).

Children with language-based learning disabilities are known to have difficulty understanding speech in background noise

(Bradlow et al., 2003; Ziegler et al., 2005). In children with dyslexia, perceptual deficits in noise occur despite normal perception in quiet conditions (Ziegler et al., 2009), indicating that the deficit may be located central to the cochlea. Consistent with the idea of centrally located noise-induced deficits, children with dyslexia can exhibit atypical cortical (Warrier et al., 2004; Wible et al., 2005) and brainstem (Cunningham et al., 2001; Russo et al., 2005) responses to speech sounds presented in white noise. However, it is not known how these noise-induced neural deficits relate to speech-in-noise (SIN) perception. In experiment 1, we examined the hypothesis that children with poor SIN perception have greater temporal delays in noise than children with good SIN perception. We predicted that children performing below the 50th percentile on a behavioral SIN task would have inordinate neural delays in multitalker babble, particularly in the region corresponding to the formant transition region of the stimulus, since this region is most perceptually vulnerable (Tallal and Stark, 1981; Banai et al., 2009; Hornickel et al., 2009). In experiment 2 using the same dataset, we examined whether children with reading impairments show degraded neural responses in noise compared with typically developing children, as predicted by the noise exclusion deficit hypothesis (Sperling et al., 2005).

Materials and Methods

Participants

Sixty-six children (ages, 8–14; mean, 10.9; SD, 1.70; 22 females) were recruited from public and private schools in the Chicago area as part of an ongoing study examining neural encoding of speech in children who are typically developing or have learning impairments. Thirty-six of these children had external diagnoses of learning impairments (29 of whom had reading impairments and 7 of whom had nonverbal learning impairments), and 30 children were normally developing. Audiometric thresholds were measured at octave intervals from 250 to 8000 Hz, and all

Received Jan. 7, 2010; revised Feb. 17, 2010; accepted Feb. 24, 2010.

This work was supported by National Institutes of Health Grant R01 DC01510 and The Hugh Knowles Center of Northwestern University. We thank Trent Nicol, Alexandra Parbery-Clark, and Jane Hornickel for their helpful comments and suggestions regarding this manuscript and Alexandra Parbery-Clark for her help with data analysis. We especially thank the children and their families who participated in the study.

Correspondence should be addressed to Dr. Nina Kraus, 2240 Campus Drive, Evanston, IL 60208. E-mail: nkraus@northwestern.edu.

DOI:10.1523/JNEUROSCI.0107-10.2010

Copyright © 2010 the authors 0270-6474/10/304922-05\$15.00/0

participants demonstrated pure-tone thresholds <20 dB with no conductive hearing loss present at two or more frequencies in either ear. Inclusionary criteria also included normal wave V click-evoked ABR latencies and normal cognitive abilities based on standard scores of ≥ 85 on verbal, performance, and overall scores of the WASI (Wechsler Abbreviated Scales of Intelligence) (Zhu and Garcia, 1999). All experimental procedures were approved by the Northwestern University Institutional Review Board.

Behavioral measures

Speech understanding in noise was evaluated with the Hearing in Noise Test (HINT) (Bio-Logic Systems), which uses the Bamford–Kowal–Bench (BKB) (Bench et al., 1979) phonetically balanced sentences appropriate for children at the first-grade reading level and above. Age-normed percentile HINT scores were used in the analysis.

To evaluate the relationship between SIN performance and literacy, reading ability was evaluated using the Test of Word Reading Efficiency–Total (TOWRE-T) (Torgesen et al., 1999), a standard test of reading efficiency. The TOWRE-T combines measures of the ability to sound out nonwords and to recognize real words quickly and accurately.

Participant groups

Top and bottom SIN groups were formed based on HINT-Front scores. The top SIN group ($N = 30$) had HINT scores ≥ 50 th percentile (mean, 78.26; SD, 15.92; range, 50–100) and the bottom SIN group ($N = 36$) had scores <50th percentile (mean, 20.28; SD, 16.20; range, 0.02–47.50). In this HINT condition, the target sentences and masking noise emanate from the same loudspeaker located 1 m directly in front of the participant. There were no significant SIN group differences for pure-tone audiometric thresholds from 250 to 8000 Hz ($p = 0.858$, independent t test; top SIN group: mean, 3.24 dB; SD, 4.33 dB; bottom SIN group: mean, 3.48 dB; SD, 4.95 dB), click-ABR latencies ($p = 0.333$, independent t test; top SIN group: mean, 5.89 ms; SD, 0.16 ms; bottom SIN group: mean, 5.84 ms; SD, 0.14 ms), or reading score ($p = 0.555$, independent t test; top SIN group: mean, 101.82; SD, 21.51; bottom SIN group: mean, 99.06; SD, 15.67).

Top and bottom reading groups were formed based on TOWRE-T reading scores and external diagnosis of reading impairment. Children in the bottom reading group ($N = 28$) had an external diagnosis of reading impairment as well as a TOWRE score <100 (TOWRE: mean, 83.5; SD, 0.944; range, 58–96), and children in the top reading group ($N = 27$) were typically developing and had a TOWRE score ≥ 100 (TOWRE: mean, 116.48; SD, 11.08; range, 101–138). There were no significant reading group differences for pure-tone audiometric thresholds from 250 to 8000 Hz ($p = 0.591$, independent t test; top readers: mean, 3.01 dB; SD, 4.49 dB; bottom readers: mean, 3.87 dB; SD, 4.78 dB), click-ABR latencies ($p = 0.921$, independent t test; top readers: mean, 5.88 ms; SD, 0.13 ms; bottom readers: mean, 5.87 ms; SD, 0.19 ms), or HINT scores ($p = 0.375$, independent t test; top readers: mean, –0.66; SD, 1.64; bottom readers: mean, –0.66; SD, 1.32).

Electrophysiology

Stimulus and recording. The speech syllable [da] was a six-formant, 170 ms syllable [described by Parbery-Clark et al. (2009b)] synthesized at a 20 kHz sampling rate using a Klatt synthesizer (Klatt, 1980). The [da] stimulus was presented with a 60 ms interstimulus interval using interleaved alternating stimulus polarities to the right ear at 80 dB sound pressure level (SPL) through an electromagnetically shielded insert earphone (ER-3; Etymotic Research) using the stimulus presentation software NeuroScan Stim2 (Sound module; Compumedics). Before each recording session, the stimulus [da] was calibrated to 80 dB SPL using a Brüel & Kjær 2238 Mediator sound level meter coupled to an insert earphone adaptor, sampling the SPL over 60 s to obtain the average SPL. The stimulus intensity was lower than levels tolerated by typically developing children and children with attention deficit disorders (Lucker et al., 1996). Responses were recorded with a vertical montage using NeuroScan Acquire4 from Cz-to-earlobe with forehead as ground at a sampling rate of 20 kHz. During electrophysiological testing, participants watched movies of their choice in a comfortable reclining chair. The left ear was unoccluded enabling the participant to hear the soundtrack played at

<40 dB SPL, an insufficient loudness to mask the stimulus. The use of movies ensured participant cooperation by enabling them to sit quietly for 2 h sessions.

The [da] was presented in two blocks: quiet (i.e., no background babble) and six-talker babble background noise. The six-talker babble (four female and two male voices) was created by mixing six tracks of sentences in Cool Edit Pro, version 2.1 (Syntrillium Software, 2003), into a 4.7 s babble track with the signal-to-noise (SNR) set at +10 dB relative to the [da] based on the root mean square (RMS) amplitude of the entire track. The Stim2 Sound program automatically tracks the level of the babble relative to the [da] syllable, keeping the SNR constant at +10 dB.

Data analysis. Electrophysiological responses were off-line bandpass filtered from 70 to 2000 Hz (12 dB/octave, zero phase-shift) to minimize low-frequency myogenic noise and cortical activity and to include energy that would be expected in the brainstem response given its phase-locking limits (Chandrasekaran and Kraus, 2010a; Skoe and Kraus, 2010). Responses were then averaged over a window of –40 to 190 ms, with time 0 corresponding to the stimulus onset. An artifact reject criterion of $\pm 35 \mu\text{V}$ was applied, and for each stimulus polarity 3000 artifact-free responses were averaged together.

The SNR of the final average response was measured by dividing the RMS of the response region (0–190 ms) of the waveform by the RMS of the prestimulus region (–40 to 0 ms). This metric was used to ensure that the response was adequately free of myogenic and electrical noise. All subjects had a minimum SNR of 1.5 in the quiet condition and 1.35 in the noise condition.

Measurement of the brainstem response

The brainstem evoked response to this 170 ms [da] syllable is characterized by three time domain regions: the onset, transition, and steady state, reflecting the corresponding characteristics of the stimulus (see Fig. 1, middle). The onset response typically has a latency of 8–11 ms and is analogous to wave V in the click response (Song et al., 2006; Chandrasekaran and Kraus, 2010a). The transition response specific to this [da] token occurs within 20–60 ms and corresponds to the consonant-to-vowel formant transition. The transition and the steady state are characterized by large, periodic peaks occurring every 10 ms, corresponding to the period of the 100 Hz fundamental frequency of the syllable.

The peaks of the brainstem responses thought to be most critical to speech perception are those that reflect important speech features, such as in the time-varying formant transition (Johnson et al., 2008; Hornickel et al., 2009). In this study, SIN- and reading-group differences were most apparent within the time-varying transition region of the response (20–60 ms). For this reason, the analysis focused on the latencies of peaks in the transition region (three positive- and negative-going peak pairs at mean latencies of ~32, 34, 42, 43, 52, and 53 ms) in both the quiet and the six-talker babble recordings (see Fig. 1, bottom). For comparison, 12 peak pairs occurring every 10 ms with the steady-state region (60–180 ms) were also evaluated. Latencies of early peaks occurring at 9 and 10 ms (defined as the response to the syllable onset) and 23 and 24 ms (defined as the response to the voicing onset) were also evaluated. Results are reported using a consistent nomenclature for the peaks (e.g., peak 32 refers to the peaks occurring at ~32 ms, etc.). The peaks were identified by the primary author and a second peak picker who was blind to group membership. In cases of disagreement over peak identification, the assistance of a third peak picker was obtained. An interpeak picker reliability measure of 92% was obtained. Although the syllable onset peaks (9 and 10 ms) and voicing onset peaks (23 and 24 ms) were consistently present in the quiet condition, their amplitudes did not exceed the noise floor in the babble condition and were therefore excluded from the analysis.

Statistical analyses

In experiment 1, we compared ABRs in the quiet and babble conditions using a two-way mixed-model multivariate analysis of covariance (ANCOVA) in SPSS with group (top SIN vs bottom SIN) serving as the between-group independent variable and condition (quiet vs noise) serving as the within-group independent variable. We covaried for reading scores to ensure that the results were not driven by reading ability, given that previous studies have demonstrated SIN deficits in children with learning impair-

ments (Bradlow et al., 2003; Ziegler et al., 2009). Positive peaks at ~32, 42, and 52 ms and negative peaks at ~34, 43, and 53 ms served as dependent variables for the analysis of the transition region of the response. In addition, positive and negative peaks from 60 to 180 ms served as dependent variables in the analysis of the steady-state region of the response.

In experiment 2, we compared ABRs in the quiet and babble conditions using a two-way mixed-model multivariate ANCOVA in SPSS with group (top readers vs bottom readers) serving as the between-group independent variable and condition (quiet vs noise) serving as the within-group independent variable. We covaried for HINT scores to ensure that the results were not driven by SIN performance. Pearson's correlations were calculated for the entire group ($N = 66$) between HINT-Front and TOWRE-T scores.

Results

For all children, background noise significantly delayed the brainstem response (Fig. 1); however, the children in the bottom SIN and bottom reading groups had greater delays in the transition period relative to the top groups. Means and SDs for each peak pair in the transition and the first two peak pairs in the steady state are provided in Table 1.

Greater timing delays in poor SIN performers (experiment 1)

A two-way mixed-model ANCOVA (including the six transition peaks) demonstrated a main effect of condition ($F_{(6,58)} = 14.984$; $p < 0.001$), indicating that noise had the expected overall effect of prolonging neural responses in both groups (Fig. 2). Furthermore, a significant interaction between SIN group and condition was noted ($F_{(6,58)} = 3.288$; $p = 0.007$). *Post hoc* analyses indicated significant group by noise effects for peaks 42 ($F_{(1,63)} = 7.879$; $p = 0.007$) and 43 ($F_{(1,63)} = 11.157$; $p = 0.001$). Figure 2, inset, demonstrates this significant interaction at one of these peaks (42), in which the peak latencies are essentially equivalent between the groups in quiet, but in noise the bottom SIN group is significantly delayed compared with the top SIN group.

A two-way mixed-model ANOVA using peaks in the steady-state portion of the response (60–180 ms) revealed no significant differences in the quiet-to-noise latency shifts between the SIN groups ($F_{(24,41)} = 1.109$; $p = 0.400$). Thus, group differences were restricted to the formant transition period in noise. Furthermore, the group differences did not result from differences in the overall magnitude of neural activity. Based on independent *t* tests, no SIN group differences were attributable to SNR ($p = 0.889$) or RMS ($p = 0.357$) differences in the quiet condition or to SNR ($p = 0.504$) or RMS ($p = 0.769$) differences in the noise condition when calculated over the entire response (0–190 ms).

Greater timing delays in poor readers (experiment 2)

A two-way mixed-model ANCOVA (including the six transition peaks) indicated a significant main effect of condition ($F_{(6,58)} = 10.611$; $p < 0.001$) as well as a significant main effect of group ($F_{(6,58)} = 2.320$; $p = 0.048$) with the bottom reading group having

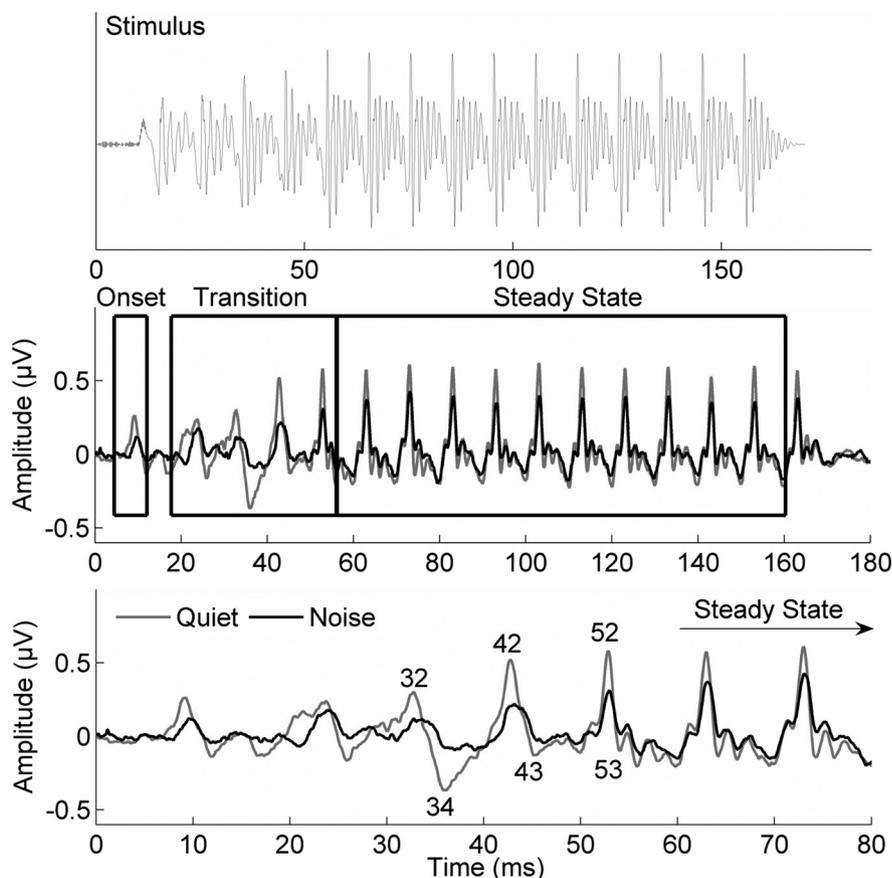


Figure 1. Top, The stimulus waveform of the speech syllable [da]. Middle, An overlay of grand average brainstem responses ($N = 66$) to the speech syllable [da] when presented in quiet (gray) and in babble noise (black). Sections of the response corresponding to the stimulus onset, formant transition, and steady-state regions are labeled. Bottom, The same response focused on the onset and transition regions where timing differences between the quiet and noise conditions are most evident. The peaks and troughs used in the analysis have been marked.

Table 1. The means and SDs are listed for eight positive- and negative-going peaks from 30 to 70 ms for top and bottom SIN and reading groups

Peak	Condition	Top SIN Mean (SD) (ms)	Bottom SIN Mean (SD) (ms)	Top readers Mean (SD) (ms)	Bottom readers Mean (SD) (ms)
32	Quiet	32.98 (0.66)	33.07 (0.60)	32.97 (0.56)	33.14 (0.74)
	Noise	34.09 (0.85)	34.06 (0.84)	33.91 (0.78)	34.11 (0.88)
34	Quiet	34.55 (0.48)	34.22 (0.66)	34.21 (0.62)	34.50 (0.56)
	Noise	35.32 (0.85)	35.39 (0.84)	35.11 (0.71)	35.53 (0.89)
42	Quiet	43.24 (0.61)	42.98 (0.42)	43.00 (0.65)	43.18 (0.58)
	Noise	43.53 (0.81)	43.89 (0.97)	43.55 (0.86)	43.83 (0.99)
43	Quiet	44.36 (0.76)	44.08 (0.46)	44.06 (0.65)	44.35 (0.65)
	Noise	44.54 (0.94)	45.05 (1.00)	44.69 (0.91)	44.82 (0.96)
52	Quiet	52.99 (0.33)	52.98 (0.37)	52.93 (0.28)	53.00 (0.43)
	Noise	53.45 (0.90)	53.57 (0.87)	53.21 (0.55)	53.76 (0.97)
53	Quiet	54.14 (0.41)	53.99 (0.33)	53.96 (0.31)	54.12 (0.44)
	Noise	54.59 (0.99)	54.65 (1.09)	54.33 (0.76)	54.84 (1.12)
63	Quiet	63.13 (0.37)	63.08 (0.28)	63.04 (0.24)	63.14 (0.40)
	Noise	63.59 (0.97)	63.52 (0.80)	63.38 (0.75)	63.58 (0.80)
64	Quiet	64.34 (0.53)	64.13 (0.25)	64.14 (0.28)	64.30 (0.54)
	Noise	64.77 (1.18)	64.60 (0.92)	64.49 (0.92)	64.71 (0.92)

greater peak timing delays than the top group (Fig. 3). *Post hoc* analyses indicated a significant group by noise effect for peak 52 ($F_{(1,63)} = 4.959$; $p = 0.030$) (Fig. 3, inset). In general, the poor readers had greater neural delays in noise. The correlation between HINT-Front and TOWRE-T ($r = 0.277$; $p = 0.024$) indicated the presence of a weak but significant relationship.

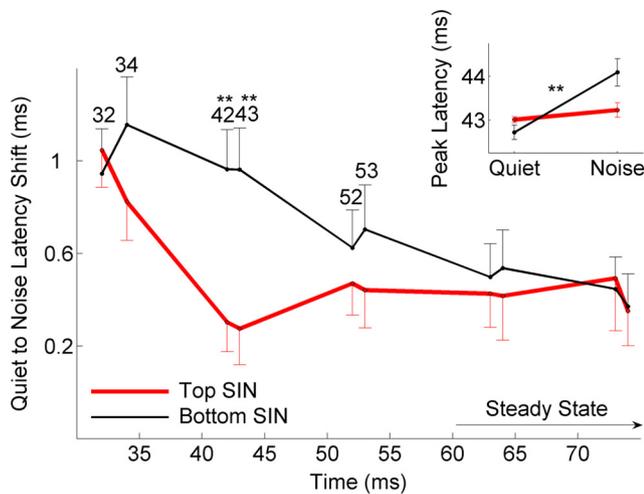


Figure 2. Timing shifts from quiet to noise for top (red heavy line) and bottom (black line) SIN groups. A significant interaction between group and condition was noted ($p < 0.01$), demonstrating greater noise-induced peak delays in the bottom SIN group. *Post hoc* analyses indicated significant group by noise effects for peaks 42 ($p < 0.01$) and 43 ($p < 0.01$). Inset, Latency interaction between the quiet and noise condition between top and bottom SIN groups for peak 42 ($p < 0.01$). The latencies are approximately equivalent in the quiet condition, but in noise the responses of the bottom group are significantly delayed. $**p < 0.01$. Error bars indicate SEM.

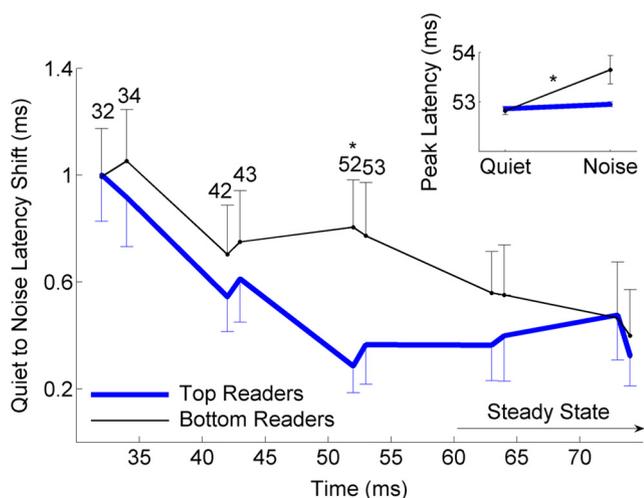


Figure 3. Timing shifts from quiet to noise for top (blue solid) and bottom (black dotted) reading groups. There was a significant main effect of group, and the bottom group had greater overall timing delays than the top reading group. Inset, Latency interaction between the quiet and noise condition between top and bottom reading groups for peak 52. $*p < 0.05$. Error bars indicate SEM.

A two-way mixed-model ANOVA using the peaks in the steady-state portion of the response (60–180 ms) revealed no significant effect of group ($F_{(24,30)} = 0.921$; $p = 0.577$). Reading group differences were not attributable to SNR ($p = 0.147$) or RMS ($p = 0.178$) differences in the quiet condition or to SNR ($p = 0.327$) or RMS ($p = 0.593$) differences in the noise condition and were therefore not a result of differences in overall neural magnitude.

Discussion

To summarize, when comparing the brainstem responses in top and bottom SIN perceivers and top and bottom readers, we found greater noise-induced timing delays in both the bottom SIN and reading groups. As we predicted, these peak delays cor-

responded to the formant transition of the evoking syllable, the most perceptually vulnerable segment of the speech syllable (Tallal and Stark, 1981). These results are in line with our prediction that deficient SIN perception and reading is associated with decreased neural synchrony leading to impaired processing of timing information in noise.

Our finding of greater noise-induced delays in bottom SIN perceivers supports the importance of temporal resolution in perception (Benasich and Tallal, 2002; Tremblay et al., 2002; Hornickel et al., 2009). The role of temporal resolution was also recently demonstrated in a study comparing brainstem timing in musicians versus nonmusicians (Parbery-Clark et al., 2009a). Musicians, who demonstrate a perceptual benefit for SIN perception (Parbery-Clark et al., 2009b), also have more robust brainstem timing in background noise relative to nonmusicians. Temporal information is important for object identification and subsequent sound segregation (Shinn-Cunningham and Best, 2008), and our results suggest that inordinate noise-induced neural delays impede the listener's ability to extract the desired signal from background noise, interfering with stream segregation at brainstem and cortical levels, ultimately leading to poorer SIN perception.

The children whose responses reflected the greatest noise-induced decreases in temporal resolution may be exhibiting the auditory analog of a noise exclusion deficit (Sperling et al., 2005). Sperling et al. showed that dyslexic and nondyslexic children performed differently on a visual task only when the visual stimulus was embedded in noise. Similarly, we found that children with poor SIN perception or reading had timing delays only in the noise condition but not in quiet, indicating that a noise exclusion deficit may also be present in the auditory system. The finding of noise-induced neural timing delays in children with either poor SIN or poor reading suggests the possibility of a common mechanism, such as a noise exclusion deficit, contributing to impairments in both of these populations.

The cause of these brainstem timing deficits is still a matter of debate. Computational models have placed the loci of SIN deficits at lower levels of the auditory system, including the brainstem, auditory nerve, and cochlea (Shamma and Klein, 2000; Carney et al., 2002). Suprathreshold temporal deficits of cochlear processing may be responsible for upstream deficits in the brainstem. Cochlear damage has also been found to degrade neural phase-locking cues in regions of audiometrically normal hearing (Lorenzi et al., 2006, 2009). However, recent work has suggested that lifelong experiences may result in top-down modulation of brainstem responses in noise (Parbery-Clark et al., 2009a). Together, this work suggests that an interplay of top-down and bottom-up processing may be needed to overcome the deleterious effects of challenging listening conditions.

This study has important implications for auditory training and other forms of intervention. Previous work has demonstrated improvement in brainstem activity after short-term auditory training in both children (Russo et al., 2005) and adults (de Boer and Thornton, 2008; Song et al., 2008) as well as improvement in cortical responses in children (Warrier et al., 2004) and adults (Tremblay et al., 2002). A number of commercially available adaptive auditory training programs (Tallal, 2004; Henderson Sabes and Sweetow, 2007; Smith et al., 2009) take advantage of exaggerated temporal cues to facilitate learning. Our results indicate that temporal training may aid SIN perception. Recent work has also demonstrated that musical training results in enhanced ability to hear speech in background noise (Parbery-Clark et al., 2009a; Chandrasekaran and Kraus, 2010b). Future and ongoing

work will help us to determine which aspects of auditory training provide the best enhancement of spectrotemporal representation and to delineate objective measures of training efficacy and predictors of success.

References

- Akhoun I, Gallégo S, Moulin A, Ménard M, Veuillet E, Berger-Vachon C, Collet L, Thai-Van H (2008) The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clin Neurophysiol* 119:922–933.
- Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Reading and subcortical auditory function. *Cereb Cortex* 19:2699–2707.
- Benasich AA, Tallal P (2002) Infant discrimination of rapid auditory cues predicts later language impairment. *Behav Brain Res* 136:31–49.
- Bench J, Kowal A, Bamford J (1979) The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Brit J Audiol* 13:108–112.
- Billings CJ, Tremblay KL, Stecker GC, Tolin WM (2009) Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hear Res* 254:15–24.
- Bradlow AR, Kraus N, Hayes E (2003) Speaking clearly for children with learning disabilities: sentence perception in noise. *J Speech Lang Hear Res* 46:80–97.
- Burkard RF, Sims D (2002) A comparison of the effects of broadband masking noise on the auditory brainstem response in young and older adults. *Am J Audiol* 11:13–22.
- Carney LH, Heinz MG, Evilsizer ME, Gilkey RH, Colburn HS (2002) Auditory phase opponency: a temporal model for masked detection at low frequencies. *Acta Acustica* 88:334–347.
- Chandrasekaran B, Kraus N (2010a) The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47:236–246.
- Chandrasekaran B, Kraus N (2010b) Music, noise-exclusion, and learning disabilities. *Music Perception* 27:297–306.
- Cunningham J, Nicol TG, Zecker SG, Bradlow AR, Kraus N (2001) Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clin Neurophysiol* 112:758–767.
- de Boer J, Thornton ARD (2008) Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J Neurosci* 28:4929–4937.
- Hall J (1992) *Handbook of auditory evoked responses*. Needham Heights, MA: Allyn and Bacon.
- Henderson Sabes J, Sweetow RW (2007) Variables predicting outcomes on Listening and Communication Enhancement (LACE) training. *Int J Audiol* 46:374–383.
- Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci U S A* 106:13022–13027.
- Johnson KL, Nicol T, Zecker SG, Bradlow AR, Skoe E, Kraus N (2008) Brainstem encoding of voiced consonant-vowel stop syllables. *Clin Neurophysiol* 119:2623–2635.
- Klatt D (1980) Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 67:971–995.
- Kraus N, Nicol TG (2005) How can the neural encoding and perception of speech be improved? In: *Plasticity and signal representation in the auditory system* (Merzenich M, Syka J, eds), pp 259–270. New York: Springer.
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A* 103:18866–18869.
- Lorenzi C, Debrulle L, Garnier S, Fleuriot P, Moore BCJ (2009) Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *J Acoust Soc Am* 125:27–30.
- Lucker JR, Geffner D, Koch W (1996) Perception of loudness in children with ADD and without ADD. *Child Psychiatry Hum Dev* 26:181–190.
- Parbery-Clark A, Skoe E, Kraus N (2009a) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci* 29:14100–14107.
- Parbery-Clark A, Skoe E, Lam C, Kraus N (2009b) Musician enhancement for speech-in-noise. *Ear Hear* 30:653–661.
- Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual organization of sound begins in the auditory periphery. *Curr Biol* 18:1124–1128.
- Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 115:2021–2030.
- Russo N, Zecker S, Trommer B, Chen J, Kraus N (2009) Effects of background noise on cortical encoding of speech in autism spectrum disorders. *J Autism Dev Disord* 39:1185–1196.
- Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N (2005) Auditory training improves neural timing in the human brainstem. *Behav Brain Res* 156:95–103.
- Shamma S, Klein D (2000) The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. *J Acoust Soc Am* 107:2631–2644.
- Shinn-Cunningham BG, Best V (2008) Selective attention in normal and impaired hearing. *Trends Amplif* 12:283–299.
- Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear*. Advance online publication. Retrieved March 9, 2010. doi:10.1097/AUD.1090b1013e3181c8db1272.
- Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, Zelinski EM (2009) A cognitive training program based on principles of brain plasticity: results from the improvement in memory with plasticity-based adaptive cognitive training (IMPACT) study. *J Am Geriatr Soc* 57:594–603.
- Song JH, Banai K, Russo NM, Kraus N (2006) On the relationship between speech- and nonspeech-evoked auditory brainstem responses. *Audiol Neurootol* 11:233–241.
- Song JH, Skoe E, Wong PC, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20:1892–1902.
- Sperling AJ, Lu ZL, Manis FR, Seidenberg MS (2005) Deficits in perceptual noise exclusion in developmental dyslexia. *Nat Neurosci* 8:862–863.
- Tallal P (2004) Improving language and literacy is a matter of time. *Nat Rev Neurosci* 5:721–728.
- Tallal P, Stark RE (1981) Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *J Acoust Soc Am* 69:568–574.
- Torgesen JK, Wagner RK, Rashotte CA (1999) *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro-Ed.
- Tremblay KL, Piskosz M, Souza P (2002) Aging alters the neural representation of speech cues. *Neuroreport* 13:1865–1870.
- Tzounopoulos T, Kraus N (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62:463–469.
- Warrier CM, Johnson KL, Hayes EA, Nicol T, Kraus N (2004) Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise. *Exp Brain Res* 157:431–441.
- Wible B, Nicol T, Kraus N (2005) Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain* 128:417–423.
- Zhu J, Garcia E (1999) *The Weschler Abbreviated Scale of Intelligence (WASI)*. New York: Psychological Corporation.
- Ziegler JC, Pech-Georgel C, George F, Alario FX, Lorenzi C (2005) Deficits in speech perception predict language learning impairment. *Proc Natl Acad Sci U S A* 102:14110–14115.
- Ziegler JS, Pech-Georgel C, George F, Lorenzi C (2009) Speech-perception-in-noise deficits in dyslexia. *Dev Sci* 12:732–745.