

# Objective Neural Indices of Speech-in-Noise Perception

Trends in Amplification  
14(2) 73–83  
© The Author(s) 2010  
Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>  
DOI: 10.1177/1084713810380227  
<http://tia.sagepub.com>



Samira Anderson, AuD<sup>1</sup> and Nina Kraus, PhD<sup>1</sup>

## Abstract

Numerous factors contribute to understanding speech in noisy listening environments. There is a clinical need for objective biological assessment of auditory factors that contribute to the ability to hear speech in noise, factors that are free from the demands of attention and memory. Subcortical processing of complex sounds such as speech (auditory brainstem responses to speech and other complex stimuli [cABRs]) reflects the integrity of auditory function. Because cABRs physically resemble the evoking acoustic stimulus, they can provide objective indices of the neural transcription of specific acoustic elements (e.g., temporal, spectral) important for hearing speech. As with brainstem responses to clicks and tones, cABRs are clinically viable in individual subjects. Subcortical transcription of complex sounds is also clinically viable because of its known experience-dependence and role in auditory learning. Together with other clinical measures, cABRs can inform the underlying biological nature of listening and language disorders, inform treatment strategies, and provide an objective index of therapeutic outcomes. In this article, the authors review recent studies demonstrating the role of subcortical speech encoding in successful speech-in-noise perception.

## Keywords

auditory brainstem, speech-in-noise perception, plasticity

## Introduction

In today's society, we are constantly bombarded by various types of background noise, and the ability to communicate in the presence of noise is an important task for successful participation in educational, social, and vocational environments. Speech-in-noise (SIN) perception is a complex task, and it poses particular demands on older adults and children with noise exclusion deficits (e.g., dyslexia, auditory processing disorders, specific language impairment, autism spectrum disorder, and attention deficit hyperactivity disorder). Auditory brainstem responses to speech and other complex stimuli (cABRs) afford the opportunity to objectively evaluate biological factors associated with successful SIN perception. In this review, we summarize studies that have examined subcortical temporal and spectral speech encoding and its relationship to SIN perception. Taken as a whole, these studies demonstrate the clinical viability of cABRs using a variety of different analysis techniques.

## Auditory Brainstem Responses to Complex Stimuli

The cABR is ideal for evaluating auditory processing because of its high degree of transparency between the stimulus waveform and the brainstem response waveform. This similarity is

apparent visually (Skoe & Kraus, 2010; Figure 1) as well as through aural demonstrations (Galbraith, Arbagey, Branski, Comerci, & Rector, 1995).<sup>1</sup> cABRs represent the features of frequency and timing in speech, music, and other stimuli, thus providing an opportunity to examine auditory processing of these behaviorally relevant sounds.

The cABR has high test–retest reliability (Russo, Nicol, Musacchia, & Kraus, 2004; Song, Nicol, & Kraus, IN PRESS), comparable with that of clicks and tone bursts (Gorga, Kaminski, Beauchaine, & Jesteadt, 1988; Hall & Mueller, 1997), and similarly, deviations on the order of fractions of milliseconds can be considered clinically significant (Banai et al., 2009; Basu, Krishnan, & Weber-Fox, 2009; Billiet & Bellis, IN PRESS; Cunningham, Nicol, Zecker, & Kraus, 2000; Wible, Nicol, & Kraus, 2004). The cABR is more effective than the click-evoked auditory brainstem response (ABR) for differentiating auditory function in typically developing children from children with auditory-based learning impairments (Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001; Song, Banai, Russo, & Kraus, 2006; Wible et al., 2004), poor SIN

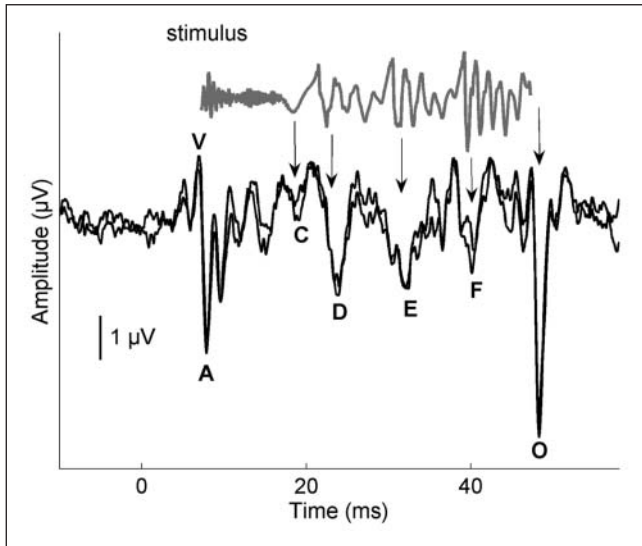
<sup>1</sup>Northwestern University, Evanston, IL, USA

### Corresponding Author:

Nina Kraus, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA

Email: [nkraus@northwestern.edu](mailto:nkraus@northwestern.edu)

Website: [www.brainvolts.northwestern.edu](http://www.brainvolts.northwestern.edu)



**Figure 1.** Time domain of a 40-ms stimulus /da/ (gray) and response (black)

*Note.* The stimulus evokes characteristic peaks in the response, labeled V, A, C, D, E, F, and O. The stimulus waveform has been shifted to account for neural lag and to allow visual alignment between peaks in the response and the stimulus. The arrows indicate where peaks in the stimulus correspond to peaks in the response. Two response waveforms of an individual participant are included to demonstrate replicability. Modified from Skoe and Kraus (2010).

abilities (Anderson, Skoe, Chandrasekaran, & Kraus, 2010; Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009), poor reading skills (Banai et al., 2005; 2009; Billiet & Bellis, IN PRESS; McAnally & Stein, 1996), poor language skills (Basu et al., 2009), and poor temporal processing abilities (Johnson, Nicol, Zecker, & Kraus, 2007).

Considerable work in our lab has examined the brainstem's response to the speech syllable /da/ because time-varying signals, in particular stop consonants, are known to be perceptually vulnerable in clinical populations (de Gelder & Vroomen, 1998; Tallal, 1980; Tallal & Stark, 1981; Tobey, Cullen, Rampp, & Fleischer-Gallagher, 1979; Townsend & Schwartz, 1981; Van Tasell, Hagen, Koblas, & Penner, 1982). The syllable consists of three time-domain components, the onset, transition, and steady state. The onset response corresponds to the onset of the consonant burst and is analogous to Wave V of the click-evoked ABR (Akhoun et al., 2008; Chandrasekaran & Kraus, 2010; Song et al., 2006). The transition region of the response, corresponding to the consonant–vowel formant transition, and the region corresponding to the steady state vowel are characterized by large periodic peaks occurring every 10 ms, paralleling the period of the 100-Hz fundamental frequency of the syllable, and smaller peaks corresponding to the harmonics.

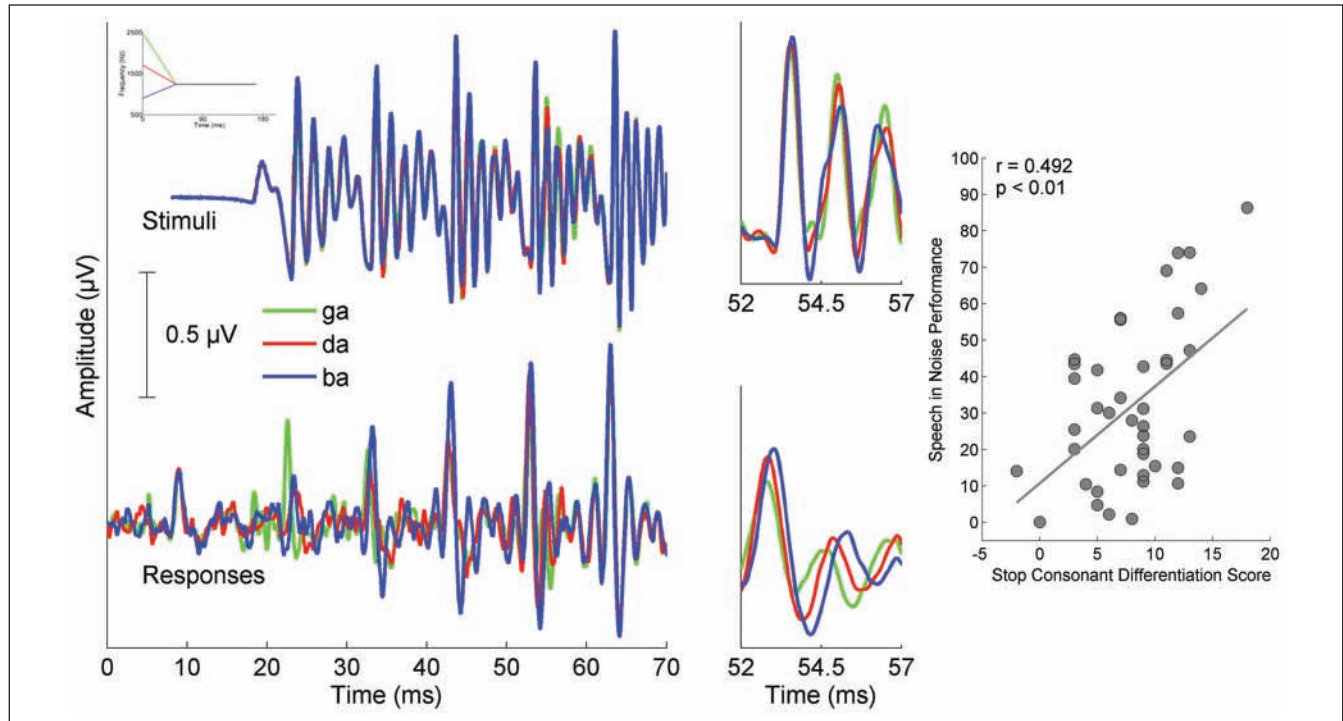
The system of afferent fibers carrying sensory information to the midbrain (inferior colliculus) and auditory cortex, and the extensive system of efferent fibers that synapse all along

the auditory pathway extending to the outer hair cells in the cochlea (Gao & Suga, 2000), support the notion that the auditory brainstem is far more than a passive conduit of information to the cortex. Indeed, the efferent fiber count may actually exceed the number of afferent fibers. The importance of the efferent pathway (from the cortex to the inferior colliculus in the brainstem) to auditory learning was demonstrated in a sound localization experiment with ferrets (Bajo, Nodal, Moore, & King, 2010). Sound localization relies on binaural cues; however, adaptation to altered cues is possible after a period of training (A. J. King et al., 2007). Bajo et al. examined the effect of pharmacological inactivation of the corticocollicular pathway on the ferret's ability to relearn localization following occlusion of one ear. They found that auditory learning in the experimental animals was significantly impaired when comparing their performance with that of controls, thus reinforcing the role of the corticocollicular pathway for auditory learning. (See Tzounopoulos & Kraus, 2009, for review of subcortical experience-dependence.)

## Speech-in-Noise Perception

Speech-in-noise perception depends on many factors involving interaction between sensory and cognitive processes. To focus on a particular speaker, the listener must form a perceptual object that enables the listener to distinguish the voice of a target speaker from other sounds (Shinn-Cunningham & Best, 2008). Object formation is determined by three primary aspects of the stimulus: location (Bronkhorst, 2000; Cherry, 1953), timing (Shinn-Cunningham & Best, 2008), and pitch (Bregman & McAdams, 1994; Brokx & Nooteboom, 1982; Darwin & Hukin, 2000; Moore, Peters, & Glasberg, 1985; Parikh & Loizou, 2005; Sayles & Winter, 2008). Pitch, derived primarily from the fundamental frequency ( $F_0$ ) and its second harmonic ( $H_2$ ) (Meddis & O'Mard, 1997), aids in voice tagging, enabling the listener to focus on the target speaker (Chandrasekaran et al., 2009).

Object formation is an important component of auditory stream segregation, or the ability to extract meaning from one particular sound source amid a background of competing sounds (Bee & Klump, 2004; Bregman, 1990; Micheyl et al., 2007; Micheyl, Tian, Carlyon, & Rauschecker, 2005; Snyder & Alain, 2007). Auditory stream segregation is mediated by attention and short-term memory (Cusack, Deeks, Aikman, & Carlyon, 2004; McLachlan & Wilson, 2010; Shinn-Cunningham & Best, 2008; Snyder, Carter, Lee, Hannon, & Alain, 2008; Sussman, Ritter, & Vaughan, 1998). For example, attention was found to modulate the streaming of alternating low- and high-frequency tones, such that the mismatch negativity response (a neural measure of auditory discrimination) was only obtained when participants were instructed to attend to high-pitched tone sequences and note deviants (low-pitched tone sequences) within this stream (Sussman et al., 1998). McLachlan and Wilson (2010) have proposed a model that engages the mechanisms of attention and short-term memory to excite an identification sequence



**Figure 2.** Stimulus timing and responses are presented for /ba/ (blue), /da/ (red), and /ga/ (green) syllables

Note. Time-domain grand average responses for 20 typically developing children (bottom panels) demonstrate timing differences in the response that reflect acoustic differences in the stimuli (top panels). The 52- to 57-ms region of the response is magnified to highlight latency differences that are present in the responses. The scatterplot on the right demonstrates a relationship between subcortical differentiation scores and speech-in-noise performance on the Hearing in Noise Test (HINT;  $r = .492$ ,  $p = .001$ ). Modified from Hornickel et al. (2009).

hierarchy, leading to modulation of thalamus and inferior colliculus spectrotemporal receptive fields to control auditory streaming.

Neurophysiologic evidence of stream segregation has been demonstrated at the level of the brainstem in the cochlear nucleus (Pressnitzer, Sayles, Micheyl, & Winter, 2008). Through single-unit recordings in the cochlear nucleus of the guinea pig, Pressnitzer et al. found evidence of stream segregation of pure tones of different frequencies presented in a repeating sequence of ABA triplets. Moreover, the neurometric functions derived from these recordings accurately predicted behavioral streaming in humans.

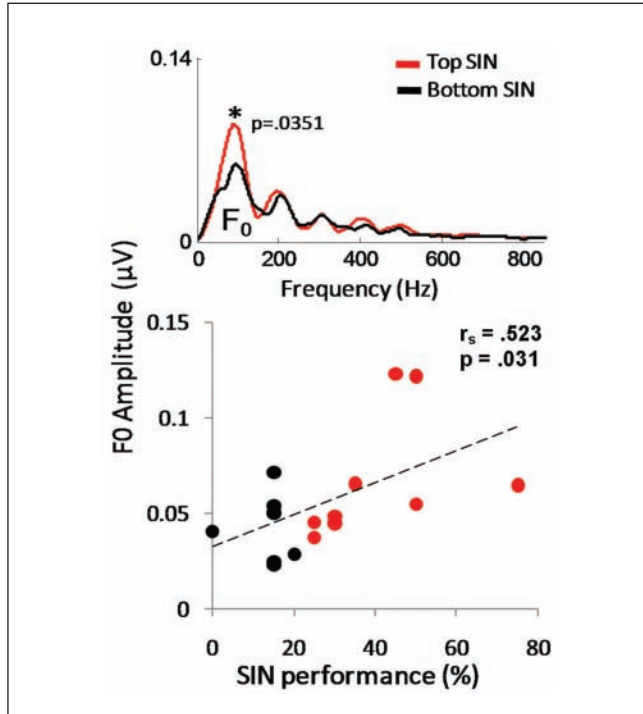
The neural mechanisms of SIN perception can also be examined in cABRs in humans. Brainstem encoding of stimulus elements (pitch, timing, and timbre) can be considered on a continuum of performance with impaired representation in poor readers (Banai et al., 2009), in children with autism spectrum disorder (Russo, Nicol, Trommer, Zecker, & Kraus, 2009), and in children with specific language impairment (Basu et al., 2009) or auditory processing disorders (Billiet & Bellis, IN PRESS) at one end of the spectrum, typically developing children in the middle of the spectrum, and auditory experts such as musicians (Bidelman, Gandour, & Krishnan, IN PRESS; Kraus & Chandrasekaran, 2010; Musacchia, Sams, Skoe, & Kraus, 2007; Parbery-Clark, Skoe, & Kraus, 2009;

Wong, Skoe, Russo, Dees, & Kraus, 2007) at the other end. Here we review subcortical encoding in children and adults who have been grouped on the basis of their SIN perception ability to improve our understanding of the biological factors contributing to SIN perception and to reach clinically viable strategies for the objective assessment of this key communication function.

### Subcortical Temporal Representation and SIN Perception Consonant Differentiation

Temporal precision is a great strength of the auditory brainstem. Stop consonants (e.g., in the syllables /ba, da, ga/) are especially vulnerable to misperception in noise (Miller & Nicely, 1955; Tallal & Stark, 1981). These syllables differ in the time-varying trajectory of the second and third formant frequencies ( $F_2$  and  $F_3$ ) during the formant transition period from the consonant to the vowel. Timing differences evoked by acoustic differences in these syllables are reflected in the cABR, with /ga/ having the shortest response latencies, /ba/ having the longest latencies, and /da/ having latencies in between the two (Johnson et al., 2008), as expected given the tonotopicity of the brainstem nuclei (Gorga et al., 1988). See Figure 2.

Comparison of response spectra to a /da/ syllable presented in background noise in top (red) and bottom (black) SIN perceivers. Group differences were found at the  $F_0$  of the stimulus (100 Hz), with the top SIN group having higher amplitudes than the bottom SIN group ( $p=.0351$ ).



**Figure 3.** The strength of  $F_0$  encoding is related to Quick Speech-in-Noise test scores (QuickSIN:  $r_s = .523$ ,  $p = .031$ ). Similar group differences for the  $F_0$  have been found in children with good and poor SIN perception (Anderson et al., IN PRESS). Adapted from Song et al., (IN PRESS).

Hornickel et al. (2009) examined the relationship between stop consonant differentiation scores and SIN perception in children with a wide range of SIN abilities. Differentiation scores were calculated for the peaks in the brainstem response, taking into account the presence of the expected timing patterns for /ba/, /da/, and /ga/, as well the magnitude of the latency differences. A relationship was found between stop consonant differentiation scores and scores on the Hearing in Noise Test (HINT; Bio-logic Systems Corp., Mundelein, IL), a behavioral measure of SIN perception (Figure 2), and children who had better SIN perception also had the expected latency patterns for the three consonant–vowel syllables ( $r = .492$ ;  $p = .001$ ). In addition, when the children were divided into thirds based on HINT performance, the top third HINT performers had significantly better stop consonant differentiation scores than the bottom HINT group.  $t(26) = 2.287$ ,  $p = .031$ .

### Noise-Induced Timing Delays

Background noise has fairly predictable effects on the ABR, including latency increases and amplitude decreases (Burkard & Sims, 2002; Russo et al., 2004). The effect of background

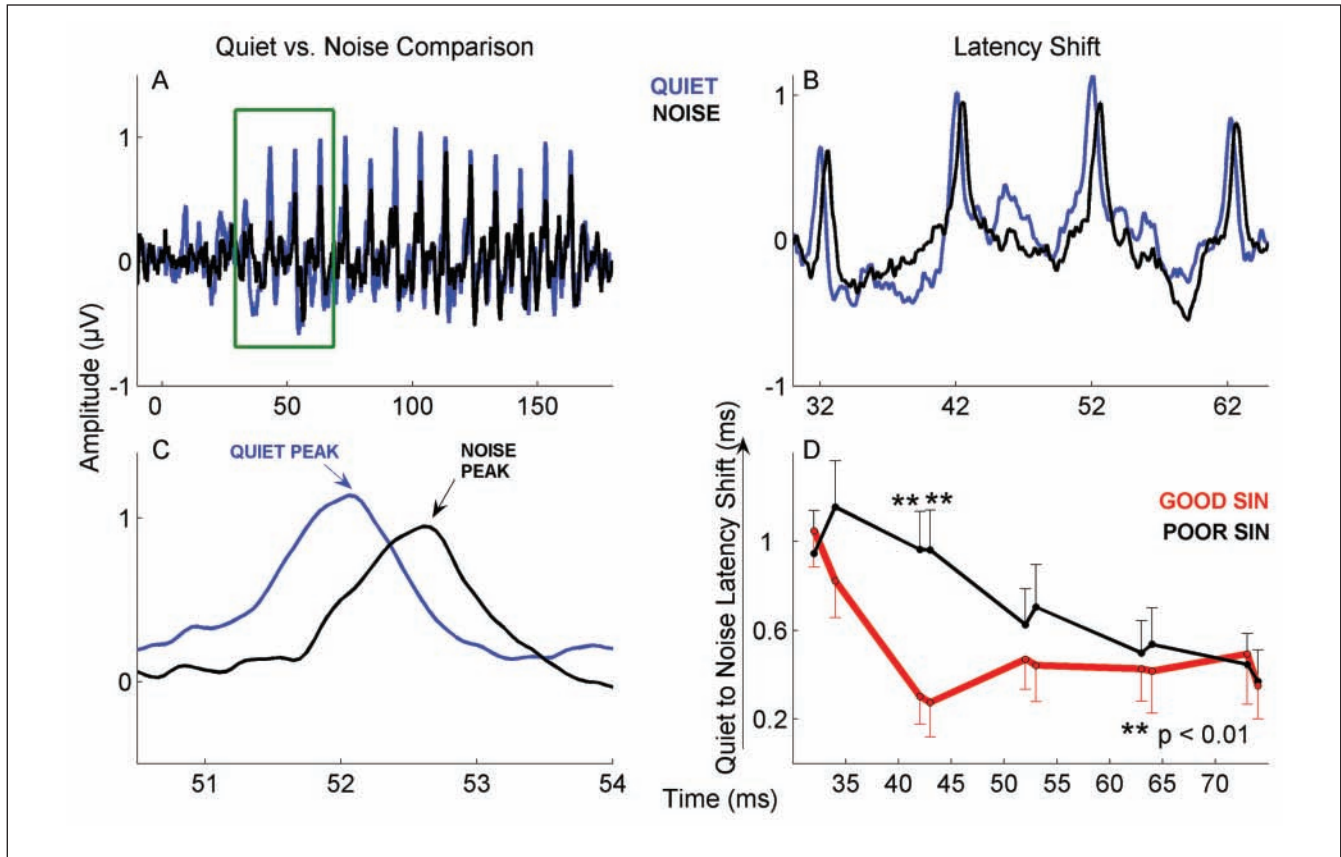
babble on the cABR was examined in children with good and poor SIN perceptions (Anderson et al., 2010). The noise effect (delay in timing) was greatest shortly after stimulus onset, delaying the response to the onset as much as a full millisecond. In children with good SIN perception (as measured by the HINT), the latency shifts leveled off by 40 ms into the response, but the shifts did not level off in children with poor SIN perception until approximately 60 ms (Figure 4). This region, which corresponds to the formant transition region of the stimulus, is the most perceptually vulnerable segment of the response (Hedrick & Younger, 2007; Náb lek, Czyzewski, & Crowley, 1994; Tallal & Stark, 1981). Overall, noise caused inordinate timing delays in the transition region of the response in children with poor SIN perception, indicating that a temporal precision deficit in the auditory brainstem is a factor in difficulties with listening in background noise.

### Subcortical Spectral Representation and SIN Perception

The role of pitch in stream segregation is well documented (Hedrick & Younger, 2007; Náb lek et al., 1994; Oxenham, 2008), and performance on segregation tasks has been shown to improve as  $F_0$  differences increase (Assmann & Summerfield, 1987; de Cheveigne, 1997; Scheffers, 1983). The strength of the  $F_0$  and lower harmonics, particularly  $H_2$ , important object-grouping cues, underlies successful perception of speech in noise (Bregman, 1990; Brokx & Nooteboom, 1982; Darwin & Hukin, 2000; Gaudrain, Grimault, Healy, & Bera, 2008; Moore et al., 1985; Parikh & Loizou, 2005; Summers & Leek, 1998; Vongpaisal, Trehub, & Schellenberg, 2006). The subcortical representation of pitch turns out to be an important factor in SIN perception, with greater subcortical representation of the  $F_0$  patterning with better SIN perception in young adults (Song, Skoe, Banai, & Kraus, IN PRESS). Speech-in-noise measures derived from the Quick Speech-in-Noise Test (QuickSIN; Etymotic Research, Elk Grove, IL; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004) correlated with the magnitude of  $F_0$  representation in response to /da/ presented in background babble ( $r_s = .523$ ;  $p = .031$ ), with larger  $F_0$  representation patterning with better SIN perception (Figure 3). This relationship was particularly pronounced in the response to the formant (time-varying) region of the syllable, the region that is considered perceptually vulnerable to the effects of background noise (Hedrick & Younger, 2007; Náb lek et al., 1994; Tallal & Stark, 1981). Notably, greater strength in the representation of pitch cues ( $F_0$  and  $H_2$ ) is linked to better speech perception ability in children (Anderson, Skoe, Chandrasekaran, Zecker, & Kraus, IN PRESS).

### Subcortical, On-Line Statistical Learning and SIN Perception

The importance of taking advantage of statistical regularities in signals is demonstrated when an infant learns to segment words from running speech. Word segmentation, an important



**Figure 4.** Effects of noise on brainstem responses in children with good and poor speech-in-noise (SIN) perceptions

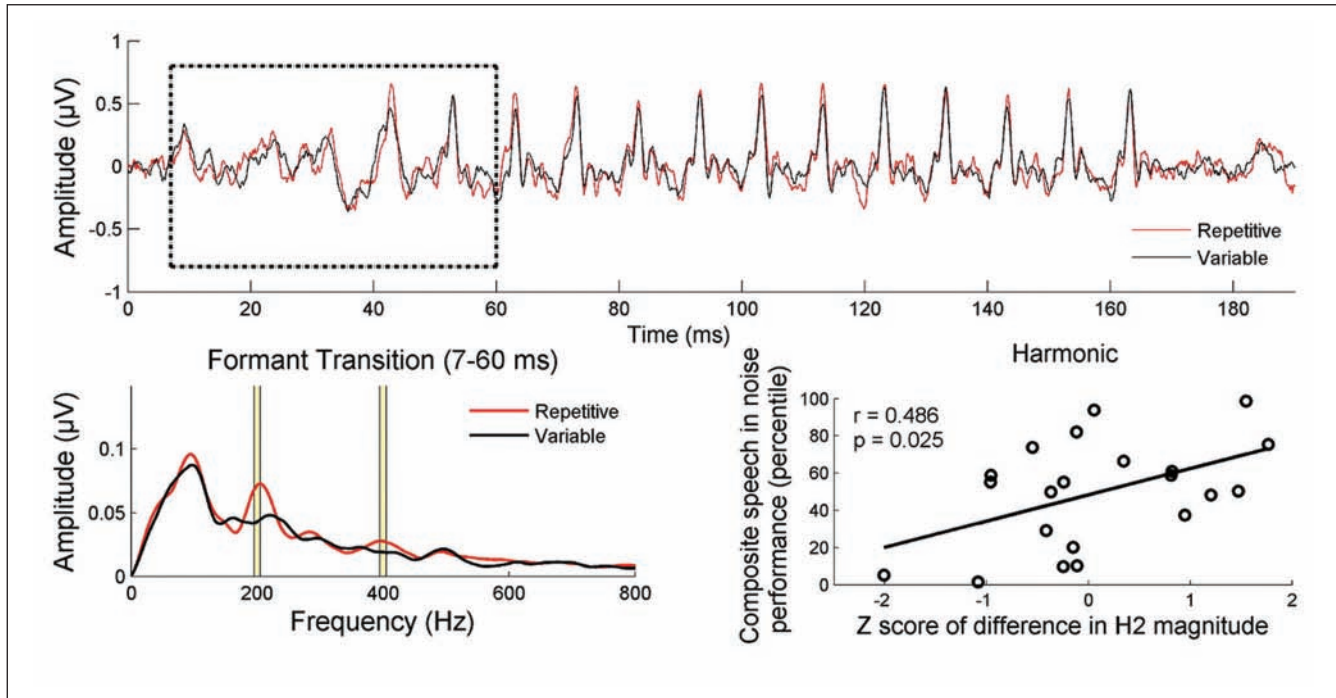
Note. The effects are most evident in the formant transition region (A, boxed) of the response from 30 to 60 ms in the grand average waveforms ( $N = 66$  children, B and C). Greater noise-induced latency shifts were noted in the children with poor speech-in-noise perception relative to children with good speech-in-noise perception ( $p < .01$ ; D). Modified from Anderson et al. (2010).

aspect of language acquisition, is based on the probability of adjacent speech sounds occurring within the same word or between words, and this statistical learning occurs remarkably quickly, within 2 minutes (Aslin, Saffran, & Newport, 1998; Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996). Adaptive (on-line) sensory processing is a hallmark of an efficient sensory system. Several studies have documented on-line plasticity in the spectrotemporal receptive fields of auditory pathway neurons (Atiani, Elhilali, David, Fritz, & Shamma, 2009; Fritz, Elhilali, & Shamma, 2005; Fritz, Elhilali, David, & Shamma, 2007; McAlpine, Martin, Mossop, & Moore, 1997). Gain and shape changes in respective fields in the ferret primary auditory cortex vary systematically with task difficulty in a tone-in-noise task, resulting in an enhanced representation of the target tone (Atiani et al., 2009). Furthermore, on-line adaptation to time-compressed speech has been observed in humans in a functional magnetic resonance imaging study, with an association between rapid task learning and increased activation of the right and left auditory association cortices and left motor cortex (Adank & Devlin, 2010).

On-line plasticity occurs in the auditory brainstem responses of children with a wide range of SIN perception abilities (Chandrasekaran et al., 2009). In typically developing children, a sharpening of brainstem responses was seen when responses to speech syllables were recorded in a predictable condition versus a variable condition (Figure 5). Importantly, the magnitude of benefit found in the regularly repeating condition correlated positively with SIN perception, as measured by the HINT. The enhancement of  $H_2$  in the predictable condition, associated with SIN perceptual ability, was interpreted as providing a mechanism for more effective “tagging” of a speaker’s voice.

### Effects of Auditory Training on Brainstem Encoding and SIN Perception

Evidence of brainstem plasticity has been demonstrated in humans. Children with learning disabilities had improved brainstem responses to speech presented in a noise background after participating in an auditory training program (Earobics;

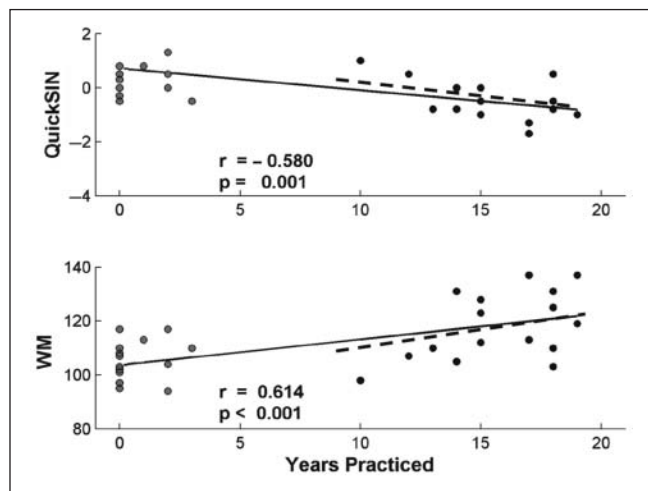


**Figure 5.** Grand average response waveforms of typically developing children ( $N = 21$ ) in response to repetitive (red) versus variable (black) presentation of speech syllable /da/ (top panel)

Note. The black rectangle outlines the formant transition region from 20 to 60 ms. Grand average spectra of repetitive versus variable presentations demonstrate enhanced magnitude of second and fourth harmonics in the repetitive condition (bottom left). The difference scores between repetitive versus variable conditions relate to Hearing in Noise Test scores ( $r = .486$ ;  $p = .025$ ; bottom right). Modified from Chandrasekaran et al. (2009).

Scientific Learning, Oakland, CA); moreover, the cABR prior to training was predictive of the benefit obtained from training (C. King, Warrier, Hayes, & Kraus, 2002; Russo, Nicol, Zecker, Hayes, & Kraus, 2005). Young adult native speakers of English showed greater accuracy in brainstem pitch tracking to a pitch contour (a nonnative speech cue) following training on a word identification task incorporating lexical pitch (Song, Skoe, Wong, & Kraus, 2008). Learning-associated brainstem changes were found in subcortical pathways revealed by otoacoustic emissions (de Boer & Thornton, 2008). These studies support the malleability of brainstem responses to complex sounds.

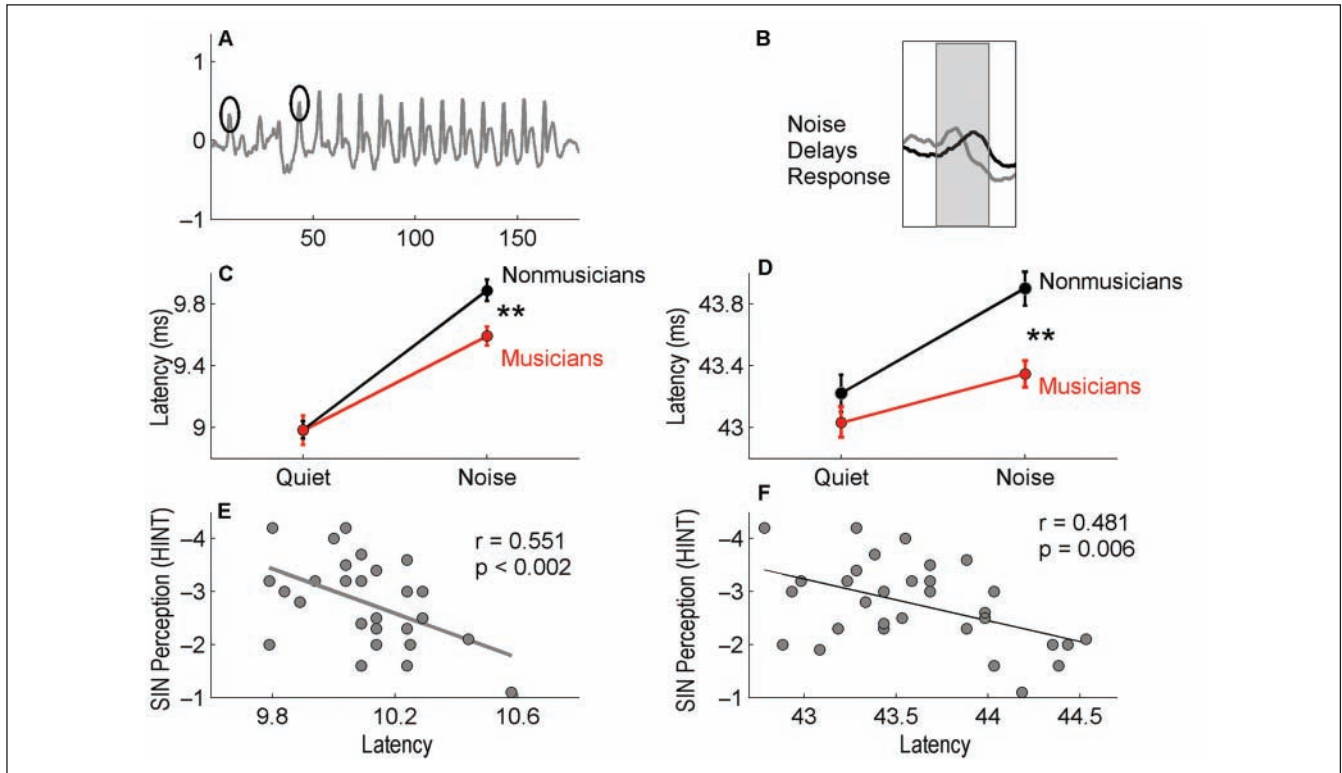
Musicians are auditory experts, and this expertise develops from years of experience selectively attending to relevant cues in a complex sound scape (e.g., melodies from background harmonics, the sound of one's own instrument). This experience can be considered a form of auditory stream segregation and object formation. Musicians have a behavioral advantage for SIN perception, having better scores on the HINT and the QuickSIN, and this benefit extends to working memory (Parbery-Clark, Skoe, Lam, & Kraus, 2009). In addition, years of consistent musical practice correlates positively both with QuickSIN and working memory (Figure 6). This study was followed by a comparison of cABRs with speech stimuli in quiet and in noise between musicians and nonmusicians. Musicians had less degradation of responses in background noise than nonmusicians (Figure 7), indicating that musical experience results in more robust subcortical speech representation



**Figure 6.** Years of musical practice relate to Quick Speech-in-Noise Test (QuickSIN) scores ( $r = -.580$ ;  $p = .001$ ) and working memory scores

Note. Composite score based on the Digits Reversed and Auditory Working Memory (WM) subtests of the Woodcock-Johnson III Test of Cognitive Abilities ( $r = .614$ ;  $p < .001$ ). Adapted from Parbery-Clark et al. (2009).

in background noise. Specifically, musicians showed less noise-induced timing delays in response to the onset and formant transition portion of the syllable /da/, consistent with parallel findings in school-age children discussed above



**Figure 7.** Comparison of brainstem responses to the speech syllable /da/ in quiet and babble noise conditions in musicians versus nonmusicians

Note. The selected peaks (onset and transition) are circled (A). Noise delays peak latencies (B), particularly in the onset and transition portions of the response. The musicians (red) show significantly shorter timing delays in noise than nonmusicians (black) for the onset (C) and transition peaks (D). The latency of the onset (E) and transition peaks (F) is correlated with SIN perception. Modified from Parbery-Clark et al. (2009).

(Anderson et al., 2010). The extent of timing disruption was correlated with SIN ability (less delay, better SIN perception on HINT measures). This enhancement of responses may be the result of top-down, corticofugal sharpening of relevant acoustic features.

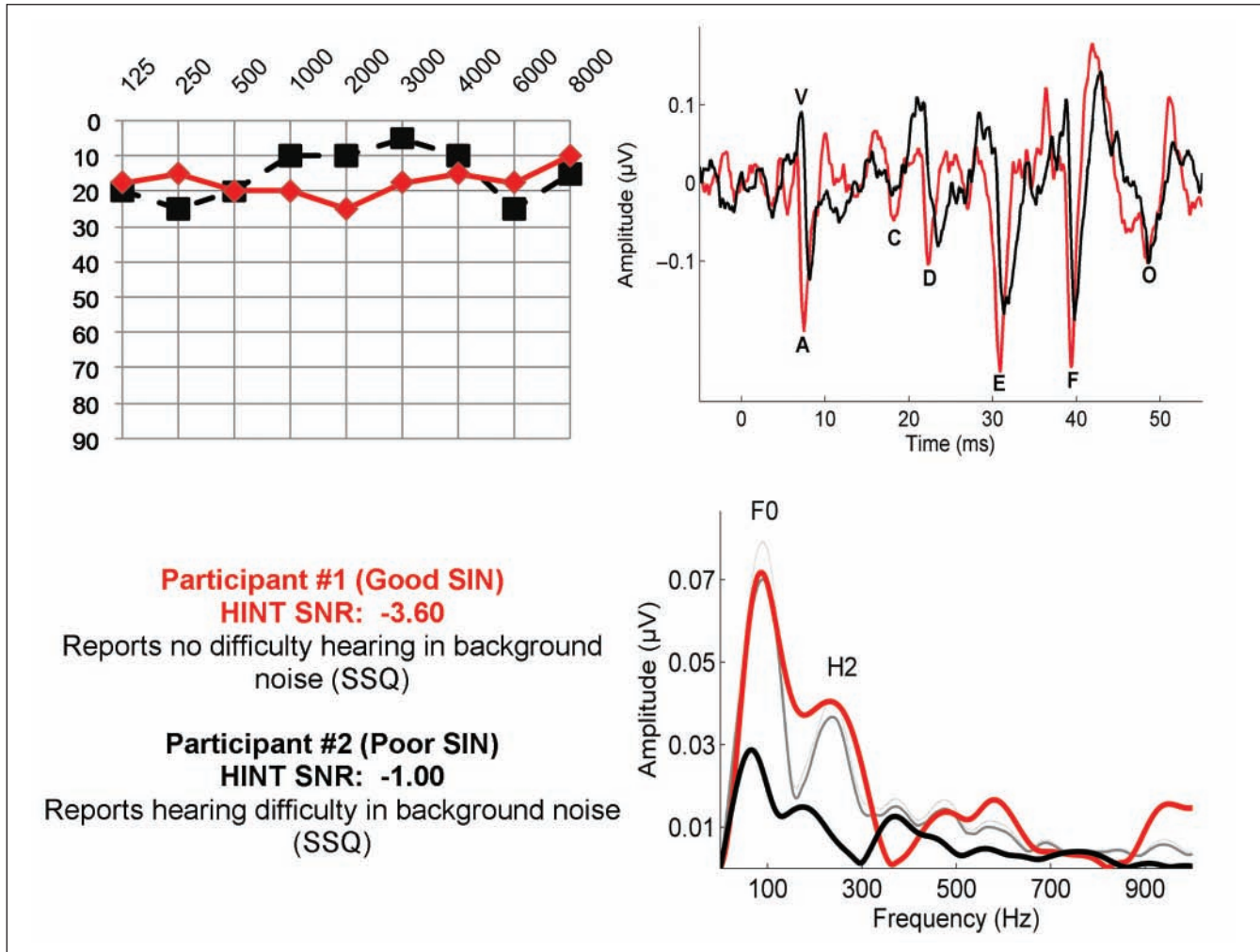
Short-term auditory training is effective in improving SIN perception. The Listening and Communication Enhancement (Neurotone, Redwood City, CA) program uses adaptive computerized auditory training to improve SIN performance in people who have hearing loss or have difficulty in hearing background noise. After using this program 30 minutes a day, 5 times a week, for 4 weeks, participants' scores on the HINT and QuickSIN improved. Improvement was also documented on self-assessment measures, including the Hearing Handicap Inventory for the Elderly/Adults (Ventry & Weinstein, 1982) and the Communication Scale for Older Adults (Kaplan, Bally, Brandt, Busacco, & Pray, 1997; see also Sweetow & Sabes, 2006).

Training with an auditory-based cognitive training program (Brain Fitness Program, Posit Science, San Francisco, CA) has revealed improvements in auditory memory and attention in adults aged 65 years and older with no history of cognitive impairment. Given the previously established relationship between auditory working memory and SIN perception

(Parbery-Clark, Skoe, Lam, & Kraus, 2009) and memory and attention and auditory stream segregation (Shinn-Cunningham & Best, 2008), this kind of training program may benefit SIN perception as well as auditory cognitive skills. We argue that training that engages both sensory and cognitive aspects of SIN perception are likely to be the most effective in driving SIN improvements (Kraus & Chandrasekaran, 2010). Germane to the topic of this review, brainstem responses can reveal this cognitive sensory interplay. Efforts are currently under way in our lab to examine the impact of auditory training on neural and behavioral measures of SIN perception.

## Case Study

Figure 8 showcases how cABR can provide an objective clinical measure of SIN perception and a metric for predicting treatment benefit and for documenting training-related changes. The figure illustrates preliminary data from an ongoing study of the relationship between subcortical processing and SIN perception. We feature two older adults who self-report good SIN perception (age 61 years) and poor SIN perception (age 62 years) and have essentially identical normal audiograms. They differ on a behavioral measure of SIN (HINT scores were  $-1.0$  vs.  $-3.6$ ) and on a subset of four



**Figure 8.** Case study demonstrating the cABRs of 2 participants with nearly identical audiograms (top left) but differing ability to hear in background noise on the Speech, Spatial, and Qualities of Hearing Scale (SSQ; 8.5/10 vs. 4/10 on four background-noise-related questions; bottom left)

Note. The participant with better SIN (Hearing in Noise Test [HINT]:  $-3.6$ , age 61 years) has faster brainstem timing (top right) and greater representation of pitch ( $F_0$  and  $H_2$ ; bottom left) than the participant with poorer SIN (HINT:  $-1.0$ , age 62 years). The solid gray line in the fast Fourier transform plot represents the grand average response of 20 normal-hearing participants ages 60 years and older, and the dotted gray line represents the standard error.

background-noise-related questions on the Speech, Spatial, and Qualities of Hearing Scale (Gatehouse & Noble, 2004; mean score of 8.25/10 vs. 4/10). We obtained brainstem responses to a 40 ms /da/ syllable using BioMARK (Biological Marker of Auditory Processing), a clinical technology available as an addition to the Navigator Pro-Auditory Evoked Potential hardware (Natus Medical Incorporated, San Carlos, CA). The individual with good SIN perception has earlier peak latencies and greater subcortical  $F_0$  representation than the individual with poor SIN perception. The differences in the brainstem responses provide quite a marked contrast and are consistent with the participants' own perception of abilities.

## Conclusions

Speech-in-noise perception is influenced by many factors, and the brainstem encoding of spectral ( $F_0$  and  $H_2$ ) and temporal features of speech appears to be a significant factor in the ability to successfully communicate in background noise. The cABR is objective, highly reliable, and can be used to examine the brainstem encoding of temporal and spectral features of speech and other complex sounds. Brainstem processing is experience-dependent, and a sharpening of responses occurs with short- and long-term training as well as on-line adaptation to stimulus regularities. The cABR continues to inform us of



the biology underlying SIN perception and promises to be a useful clinical tool in the assessment and management of SIN perception difficulties.

### Acknowledgments

We would especially like to thank the members of the Auditory Neuroscience Laboratory and our participants for their contributions to this work.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article:

This work was funded by the National Institutes of Health (RO1 DC01510) and the National Science Foundation SGER (0842376).

### Note

1. See demonstration on home page: <http://www.brainvolts.northwestern.edu>.

### References

- Adank, P., & Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage*, *49*, 1124-1132.
- Akhoun, I., Gallégo, S., Moulin, A., Ménard, M., Veillet, 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. *Clinical Neurophysiology*, *119*, 922-933.
- Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *Journal of Neuroscience*, *30*, 4922-4926.
- Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (IN PRESS). Brainstem correlates of speech-in-noise perception in children. *Hearing Research*.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321-324.
- Assmann, P. F., & Summerfield, Q. (1987). Perceptual segregation of concurrent vowels. *Journal of the Acoustical Society of America*, *82*, S120.
- Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, *61*, 467-480.
- Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, *13*, 253-260.
- Banai, K., Nicol, T., Zecker, S., & Kraus, N. (2005). Brainstem timing: Implications for cortical processing and literacy. *Journal of Neuroscience*, *25*, 9850-9857.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex*, *19*, 2699-2707.
- Basu, M., Krishnan, A., & Weber-Fox, C. (2009). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Developmental Science*, *13*, 77-91.
- Bee, M. A., & Klump, G. M. (2004). Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology*, *92*, 1088-1104.
- Bidelman, G., Gandour, J., & Krishnan, A. (IN PRESS). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*. doi:10.1162/jocn.2009.21362
- Billiet, C., & Bellis, T. J. (IN PRESS). The relationship between brainstem temporal processing and performance on tests of central auditory function in children with reading disorders. *Journal of Speech, Language, and Hearing Research*.
- Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge: MIT Press.
- Bregman, A. S., & McAdams, S. (1994). Auditory scene analysis: The perceptual organization of sound. *Journal of the Acoustical Society of America*, *95*, 1177-1178.
- Brox, J. P., & Nooteboom, S. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, *10*, 23-26.
- Bronkhorst, A. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica*, *86*, 117-128.
- Burkard, R. F., & Sims, D. (2002). A comparison of the effects of broadband masking noise on the auditory brainstem response in young and older adults. *American Journal of Audiology*, *11*, 13-22.
- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T. G., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron*, *64*, 311-319.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*, 236-246.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975-979.
- Cunningham, J., Nicol, T. G., Zecker, S. G., Bradlow, A. R., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology*, *112*, 758-767.
- Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: Development and behavioral correlates of perception. *Ear and Hearing*, *21*, 554-568.
- Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 643-656.
- Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *Journal of the Acoustical Society of America*, *107*, 970-977.

- de Boer, J., & Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience*, 28, 4929-4937.
- de Cheveigne, A. (1997). Concurrent vowel identification. III. A neural model of harmonic interference cancellation. *Journal of the Acoustical Society of America*, 101, 2857-2865.
- de Gelder, B., & Vroomen, J. (1998). Impaired speech perception in poor readers: Evidence from hearing and speech reading. *Brain & Language*, 64, 269-281.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research*, 229, 186-203.
- Fritz, J., Elhilali, M., & Shamma, S. (2005). Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing Research*, 206, 159-176.
- Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport*, 6, 2363-2367.
- Gao, E., & Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: Role of the corticofugal system. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 8081-8086.
- Gatehouse, S., & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology*, 43, 85-99.
- Gaudrain, E., Grimault, N., Healy, E. W., & Bera, J.-C. (2008). Streaming of vowel sequences based on fundamental frequency in a cochlear-implant simulation. *Journal of the Acoustical Society of America*, 124, 3076-3087.
- Gorga, M. P., Kaminski, J. R., Beauchaine, K. A., & Jesteadt, W. (1988). Auditory brainstem responses to tone bursts in normally hearing subjects. *Journal of Speech, Language, and Hearing Research*, 31, 87-97.
- Hall, J. W., & Mueller, H. G. (1997). *Audiologists' desk reference* (Vol. 1). San Diego, CA: Singular.
- Hedrick, M. S., & Younger, M. S. (2007). Perceptual weighting of stop consonant cues by normal and impaired listeners in reverberation versus noise. *Journal of Speech, Language and Hearing Research*, 50, 254-269.
- Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 13022-13027.
- Johnson, K. L., Nicol, T., Zecker, S. G., Bradlow, A. R., Skoe, E., & Kraus, N. (2008). Brainstem encoding of voiced consonant-vowel stop syllables. *Clinical Neurophysiology*, 119, 2623-2635.
- Johnson, K., Nicol, T. G., Zecker, S. G., & Kraus, N. (2007). Auditory brainstem correlates of perceptual timing deficits. *Journal of Cognitive Neuroscience*, 19, 376-385.
- Kaplan, H., Bally, S., Brandt, F., Busacco, D., & Pray, J. (1997). Communication scale for older adults (CSOA). *Journal of the American Academy of Audiology*, 8, 203-217.
- Killion, M., Niquette, P., Gudmundsen, G., Revit, L., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 116, 2395-2405.
- King, A. J., Bajo, V. M., Bizley, J. K., Campbell, R. A. A., Nodal, F. R., Schulz, A. L., & Schnupp, J. W. H. (2007). Physiological and behavioral studies of spatial coding in the auditory cortex. *Hearing Research*, 229, 106-115.
- King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters*, 319, 111-115.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35-B42.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11, 599-605.
- McAlpine, D., Martin, R., Mossop, J., & Moore, D. (1997). Response properties of neurons in the inferior colliculus of the monaurally deafened ferret to acoustic stimulation of the intact ear. *Journal of Neurophysiology*, 78, 767-779.
- McAnally, K., & Stein, J. (1996). Auditory temporal coding in dyslexia. *Proceedings: Biological Sciences*, 263, 961-965.
- McLachlan, N., & Wilson, S. (2010). The central role of recognition in auditory perception: A neurobiological model. *Psychological Review*, 117, 175-196.
- Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. *Journal of the Acoustical Society of America*, 102, 1811-1820.
- Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A. J., Rauschecker, J. P., . . . Wilson, E. C. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Research*, 229, 116-131.
- Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron*, 48, 139-148.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338-352.
- Moore, B. C., Peters, R. W., & Glasberg, B. R. (1985). Thresholds for the detection of inharmonicity in complex tones. *Journal of the Acoustical Society of America*, 77, 1861-1867.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, 104, 15894-15898.
- Náb lek, A. K., Czyzewski, Z., & Crowley, H. (1994). Cues for perception of the diphthong /i/ in either noise or reverberation. Part I. Duration of the transition. *Journal of the Acoustical Society of America*, 95, 2681-2693.
- Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, 12, 316-331.

- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, *29*, 14100-14107.
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*, 653-661.
- Parikh, G., & Loizou, P. C. (2005). The influence of noise on vowel and consonant cues. *Journal of the Acoustical Society of America*, *118*, 3874-3888.
- Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology*, *18*, 1124-1128.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, *115*, 2021-2030.
- Russo, N., Nicol, T., Trommer, B., Zecker, S., & Kraus, N. (2009). Brainstem transcription of speech is disrupted in children with autism spectrum disorders. *Developmental Science*, *12*, 557-567.
- Russo, N., Nicol, T., Zecker, S., Hayes, E., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, *156*, 95-103.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- Sayles, M., & Winter, I. M. (2008). Ambiguous pitch and the temporal representation of inharmonic iterated rippled noise in the ventral cochlear nucleus. *Journal of Neuroscience*, *28*, 11925-11938.
- Scheffers, M. T. M. (1983). Simulation of auditory analysis of pitch: An elaboration on the DWS pitch meter. *Journal of the Acoustical Society of America*, *74*, 1716-1725.
- Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*, 283-299.
- Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing*, *31*, 302-324.
- Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin*, *133*, 780-799.
- Snyder, J. S., Carter, O. L., Lee, S.-K., Hannon, E. E., & Alain, C. (2008). Effects of context on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1007-1016.
- Song, J. H., Banai, K., Russo, N. M., & Kraus, N. (2006). On the relationship between speech- and nonspeech-evoked auditory brainstem responses. *Audiology & Neurotology*, *11*, 233-241.
- Song, J. H., Nicol, T., & Kraus, N. (IN PRESS). Test-retest reliability of the speech-evoked auditory brainstem response in young adults. *Clinical Neurophysiology*.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2009, October). *Enhancement of brainstem encoding of the fundamental frequency in listeners with good speech perception in noise*. Poster presented at the Society for Neuroscience Annual Meeting, Chicago, IL.
- Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, *20*, 1892-1902.
- Summers, V., & Leek, M. R. (1998). F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. *Journal of Speech, Language and Hearing Research*, *41*, 1294-1306.
- Sussman, E., Ritter, W., & Vaughan, H. G. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research*, *789*, 130-138.
- Sweetow, R. W., & Sabes, J. H. (2006). The need for and development of an adaptive listening and communication enhancement (LACE) program. *Journal of the American Academy of Audiology*, *17*, 538-558.
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain & Language*, *9*, 182-198.
- Tallal, P., & Stark, R. E. (1981). Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America*, *69*, 568-574.
- Tobey, E. A., Cullen, J. K., Jr., Rampp, D. L., & Fleischer-Gallagher, A. M. (1979). Effects of stimulus-onset asynchrony on the dichotic performance of children with auditory-processing disorders. *Journal of Speech, Language, and Hearing Research*, *22*, 197-211.
- Townsend, T. H., & Schwartz, D. M. (1981). Error analysis on the California consonant test by manner of articulation. *Ear and Hearing*, *2*, 108-111.
- Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*, 463-469.
- Van Tasell, D., Hagen, L. T., Koblas, L. L., & Penner, S. G. (1982). Perception of short-term spectral cues for stop consonant place by normal and hearing-impaired subjects. *Journal of the Acoustical Society of America*, *72*, 1771-1780.
- Ventry, I., & Weinstein, B. (1982). The hearing handicap inventory for the elderly: A new tool. *Ear and Hearing*, *3*, 128-134.
- Vongpaisal, T., Trehub, S. E., & Schellenberg, E. G. (2006). Song recognition by children and adolescents with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *49*, 1091-1103.
- Wible, B., Nicol, T., & Kraus, N. (2004). Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychology*, *67*, 299-317.
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420-422.