Sensory-Cognitive Interaction in the Neural Encoding of Speech in Noise: A Review

Samira Anderson*†
Nina Kraus*†‡§

Abstract

Background: Speech-in-noise (SIN) perception is one of the most complex tasks faced by listeners on a daily basis. Although listening in noise presents challenges for all listeners, background noise inordinately affects speech perception in older adults and in children with learning disabilities. Hearing thresholds are an important factor in SIN perception, but they are not the only factor. For successful comprehension, the listener must perceive and attend to relevant speech features, such as the pitch, timing, and timbre of the target speaker's voice. Here, we review recent studies linking SIN and brainstem processing of speech sounds.

Purpose: To review recent work that has examined the ability of the auditory brainstem response to complex sounds (cABR), which reflects the nervous system's transcription of pitch, timing, and timbre, to be used as an objective neural index for hearing-in-noise abilities.

Study Sample: We examined speech-evoked brainstem responses in a variety of populations, including children who are typically developing, children with language-based learning impairment, young adults, older adults, and auditory experts (i.e., musicians).

Data Collection and Analysis: In a number of studies, we recorded brainstem responses in quiet and babble noise conditions to the speech syllable /da/ in all age groups, as well as in a variable condition in children in which /da/ was presented in the context of seven other speech sounds. We also measured speech-in-noise perception using the Hearing-in-Noise Test (HINT) and the Quick Speech-in-Noise Test (QuickSIN).

Results: Children and adults with poor SIN perception have deficits in the subcortical spectrotemporal representation of speech, including low-frequency spectral magnitudes and the timing of transient response peaks. Furthermore, auditory expertise, as engendered by musical training, provides both behavioral and neural advantages for processing speech in noise.

Conclusions: These results have implications for future assessment and management strategies for young and old populations whose primary complaint is difficulty hearing in background noise. The cABR provides a clinically applicable metric for objective assessment of individuals with SIN deficits, for determination of the biologic nature of disorders affecting SIN perception, for evaluation of appropriate hearing aid algorithms, and for monitoring the efficacy of auditory remediation and training.

Key Words: Auditory brainstem response, evoked potentials, frequency, musicians, speech in noise, timing

Abbreviations: ABR = auditory brainstem response; cABR = auditory brainstem response to complex sounds; FFR = frequency following response; HINT = Hearing-in-Noise Test; QuickSIN = Quick Speech-in-Noise Test; SIN = speech-in-noise

*Auditory Neuroscience Laboratory, Northwestern University; †Department of Communication Sciences, Northwestern University; ‡Department of Neurobiology and Physiology, Northwestern University; §Department of Otolaryngology, Northwestern University

Nina Kraus, PhD, 2240 Campus Dr., Evanston, IL 60208; Phone: 847-491-3181; Fax: 847-467-1464; E-mail: nkraus@northwestern.edu; Web site: www.brainvolts.northwestern

575

INTRODUCTION

Most listening environments are filled with various types of background noise, and the most troubling noise is often the competing speech heard in restaurants, school cafeterias, and classrooms. Successful speech-in-noise (SIN) perception is a vital part of everyday life, enabling listeners to participate in social, vocational, and educational activities. Children, especially those with learning disabilities, and older adults are particularly vulnerable to the effects of noise on speech perception (Bradlow et al, 2003; Ziegler et al, 2005; Kim et al, 2006). These difficulties may be seen in the presence of audiometrically normal hearing, suggesting that deficits central to the cochlea may be a factor in SIN perception (Humes, 1996; Kim et al, 2006). It has been proposed that some learning disabilities in children may result in part from a noise exclusion deficit, which manifests in the presence of noise but not in quiet situations (Sperling et al, 2005; Ziegler et al, 2009). In older adults, impaired perception may result from age-related factors affecting neural synchrony (Frisina and Frisina, 1997; Schneider and Pichora-Fuller, 2001; Tremblay et al, 2003; Caspary et al, 2005).

SIN perception is a complex task involving interplay of sensory and cognitive processes. In order to identify the target sound or speaker from a background of other noises, the listener must first form an auditory object based on spectrotemporal cues (Bronkhorst, 2000; Best et al, 2007; Shinn-Cunningham and Best, 2008). Object formation is a necessary step in stream segregation, a process that allows the listener to extract meaning from an auditory environment filled with multiple sound sources (Bregman, 1990; Bee and Klump, 2004; Micheyl et al, 2007; Snyder and Alain, 2007). Vocal pitch, as defined largely by the fundamental frequency (F_0) and the second harmonic of the stimulus (H_2), is important for auditory grouping, allowing the listener to "tag" or attach a particular identity to the speaker's voice (Brokx and Nooteboom, 1982; Moore et al, 1985; Bregman and McAdams, 1994; Darwin and Hukin, 2000; Parikh and Loizou, 2005; Sayles and Winter, 2008). The ability to form auditory objects and to segregate multiple sound sources into distinct streams is mediated, at least in part, by top-down cognitive processes such as attention and short-term memory (Best et al, 2007; Heinrich et al, 2007).

The characteristics of the speech signal that make it possible to extract the target speech from competing background noise include pitch (F_0), timing (speech onsets, offsets, and transitions between phonemes), and timbre (harmonics). These aspects of speech are well represented in the auditory brainstem response to complex sounds (cABR). The frequency following response (FFR) of the cABR is well-suited for the evaluation of the centrally based processes involved in SIN perception as it mimics the sound input remarkably well both in the time and frequency domains (Galbraith et al, 1995), and it is reliable and consistent across time (Kraus and Nicol, 2005; Song, Nicol, et al, in press) (Fig. 1). The auditory brainstem response (ABR) to a consonant-vowel syllable (e.g., /da/) is characterized by three time-domain regions: the onset, transition, and steady state, reflecting the corresponding characteristics of the stimulus. The onset response is analogous to wave V in the click response (Song et al, 2006; Chandrasekaran and Kraus, 2010b). The transition response specific to this /da/ token corresponds to the consonant-to-vowel formant transition. The transition and the steady state are characterized by large, periodic peaks occurring every 10 msec, corresponding to the period of the 100 Hz fundamental frequency of the syllable. The neural phase locking activity underlying the FFR represents the periodicity of the stimulus up to about 1500 Hz, the phase locking limit of the brainstem (Chandrasekaran and Kraus, 2010b).

Like the click-evoked response, peak latency differences on the order of fractions of milliseconds can be clinically significant in the cABR (Wible et al, 2004; Chandrasekaran and Kraus, 2010b). Furthermore, the cABR is experience dependent, and changes in the response have been demonstrated as the result of short-term auditory training and life-long experiences with language and music (Krishnan et al, 2005; Song, Skoe, et al, 2008; Tzounopoulos and Kraus, 2009; Kraus and Chandrasekaran, 2010) and online tracking of stimulus regularities (Chandrasekaran, Hornickel, et al, 2009).

Spectrotemporal features of the cABR relate to cognitive processes such as language (Banai et al, 2005; Krishnan et al, 2005) and music (Musacchia et al, 2007; Parbery-Clark, Skoe, Kraus, 2009; Strait et al, 2009b), thus providing a mechanism for the evaluation of cognitive influences on lower-level auditory function. It is thought that auditory brainstem function is modulated by higher-level processes via top-down processing. This cognitive-sensory interaction is made possible by a multitude of afferent fibers carrying sensory information to the midbrain (inferior colliculus) and auditory cortex in concert with the corticofugal pathway, an extensive system of descending efferent fibers that synapse all along the auditory pathway, extending even to the outer hair cells of the basilar membrane (Gao and Suga, 2000).

A number of different approaches have been used to examine brainstem encoding of speech syllables, including the measurement of frequency and timing information. Effort has also been made to quantify the auditory brainstem's ability to profit from regularities in an ongoing speech stream. Here we review several studies performed in the Auditory Neuroscience Laboratory at Northwestern University that link
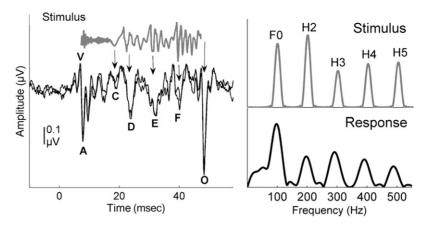
Figure 1. In the left panel, the time domains of a 40 msec stimulus /da/ (gray) and auditory brainstem response (black) are pictured. The stimulus evokes characteristic peaks in the response, labeled as *V, A, C, D, E, F,* and *O*. The stimulus waveform has been shifted to account for neural lag and to allow visual alignment between peaks in the response and the stimulus, which are indicated by arrows. Two responses from the same individual are shown to demonstrate replicability. In the right panel are the spectra of the stimulus and response. Adapted from Skoe and Kraus, 2010.

auditory brainstem encoding of speech with SIN perception across populations.

The Role of Brainstem Pitch Encoding and SIN Perception

Studies with children (Anderson, Skoe, Chandrasekaran, Zecker, et al, 2010), young adults (Song, Skoe, et al, in press), and older adults, including those with normal hearing and mild hearing impairment (Anderson et al, 2009), have examined the role that the auditory brainstem encoding of low frequencies (F_0 and H_2) plays in SIN perception. The lower harmonics are essential acoustic contributors to pitch perception (Meddis and O'Mard, 1997), and pitch cues aid in object formation and the ability to "tag" a speaker's voice (Oxenham, 2008; Shinn-Cunningham and Best, 2008; Chandrasekaran, Hornickel, et al, 2009). In a recent study, children ages 8 to 14 were divided into groups of good and poor SIN perception based on percentile scores on the HINT (Hearing-in-Noise Test; Natus Medical, Inc., San Carlos, CA) (Anderson, Skoe, Chandrasekaran, Zecker, et al, 2010). Brainstem responses were recorded to the speech syllable /da/ without competing background noise, and fast Fourier transforms (FFTs) were calculated for the transition regions of the response (20–60 msec) using 100 Hz bins centered around the F_0 of 100 Hz and its integer multiples. Added alternating polarities, emphasizing the envelope of the response and F_0, were used in this study (Aiken and Picton, 2008; Skoe and Kraus, 2010). The good SIN perceivers had greater spectral magnitudes for the F_0 and H_2 compared to poor SIN perceivers. Therefore, just as behavioral studies have revealed the importance of pitch for object identification and stream segregation, this study demonstrated that the robustness of subcortical encoding of pitch (F_0 and H_2) is a significant factor in

SIN perception. Greater representation of these low frequencies indicates better phase locking and neural synchrony, which results in the response being more resistive to the degradative effects of noise.

The feasibility of assessing pitch processing in the brainstem has been demonstrated in young adults (Jeng et al, 2010). Furthermore, robust subcortical encoding of pitch is important for hearing speech in noise in young (Song, Skoe, et al, in press) and older adults (Anderson et al, 2009) as well as children (Anderson, Skoe, Chandrasekaran, Kraus, 2010). Young adults were divided into two groups of top and bottom SIN performers based on scores on the Quick Speech-in-Noise Test (Quick-SIN) (Etymotic Research; Killion et al, 2004). Brainstem responses were recorded in these participants to the speech syllable /da/ when presented in a background of six-talker babble. The importance of F_0 encoding at the level of the auditory brainstem was noted in the FFTs, which indicated that good SIN perceivers have stronger F_0 activation in noise than poor SIN perceivers. Finally, in a study with older adults, F_0 magnitudes of brainstem responses in noise were significantly higher in good than in poor SIN perceivers (based on HINT scores). Taken together, these studies demonstrate that auditory brainstem representation of the F_0 and H_2 correlate with SIN perception across the age span (school-age children to older adults).

Utilizing Stimulus Regularities and SIN Perception

How the auditory brainstem makes use of stimulus regularities is important for forming a perceptual anchor in order to extract the desired talker's voice from a background of competing voices. A perceptual anchor is a type of memory trace that links perception with memory (Ahissar et al, 2006), and it is formed in response to
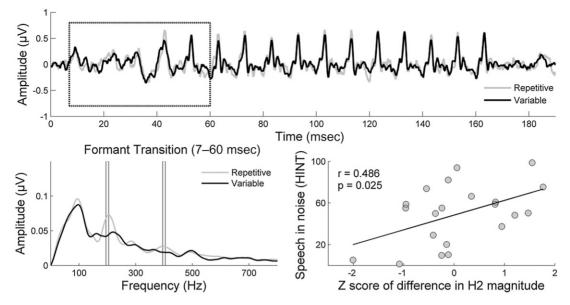
Figure 2. Grand average response waveforms of typically developing children (N = 21) in response to repetitive (gray) versus variable (black) presentation of a 170 msec speech syllable /da/ *(top panel)*. Brainstem responses in regularly occurring (gray) versus variable (black) presentations of the /da/ syllable differ in their frequency spectra, with enhanced representation of H2 and H4 (over 10 Hz bins represented by vertical lines) noted in the regular presentation *(bottom left)*. The differences in spectral amplitude of H2 and H4 (7–60 msec) between the two conditions (repetitive context minus variable context) were calculated for each child and normalized to the group mean by converting to a z-score. The normalized difference in H2 magnitude between the regularly occurring and variable conditions is related to SIN performance as measured by the Hearing-in-Noise Test (HINT) *(bottom right)*. Adapted from Chandrasekaran, Hornickel, et al, 2009.

regularly repeating stimuli. Perceptual anchors enable typically developing children to make the comparative discriminations needed when listening in background noise (Ahissar et al, 2006; Ahissar, 2007; Chandrasekaran, Hornickel, et al, 2009). Ahissar et al (2006) compared SIN performance in typically developing children with dyslexia using sets of either 10 or 40 pseudowords. They found that the children with dyslexia experienced performance deficits only with the small set of 10 stimuli, and they reasoned that the superior performance of the typically developing children was due in part to their ability to profit from stimulus repetition in order to improve performance. Our laboratory evaluated auditory brainstem adaptation to regularities in predictable versus variable speech streams in typically developing children, for which we hypothesized an auditory brainstem enhancement of predictable stimuli related to the formation of perceptual anchors (Chandrasekaran, Hornickel, et al, 2009). Auditory brainstem function in typically developing children was compared to children with developmental dyslexia in a paradigm similar to that of Ahissar's 2006 study. When comparing auditory brainstem responses to the speech syllable /da/ presented in a predictable context (in which the /da/ is the only syllable presented) to responses recorded in a variable context (in which the /da/ is presented randomly amid seven other speech syllables), greater H_2 and H_4 amplitudes were found in responses to the predictable condition in typically developing children. Despite the large response

variability, the degree of amplitude difference between these two conditions correlated with SIN perception as measured by the HINT (Fig. 2). Children with dyslexia were unable to benefit from stimulus regularities, as indicated by the lack of difference between the regularly repeating and variable presentations. These results indicate that
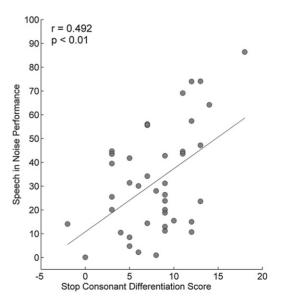

Figure 3. Subcortical differentiation of stop consonants (/ba/, /da/, and /ga/) is related to SIN performance on the HINT. Children with better subcortical differentiation scores have higher HINT scores ($p < 0.01$). Adapted from Hornickel et al, 2009.
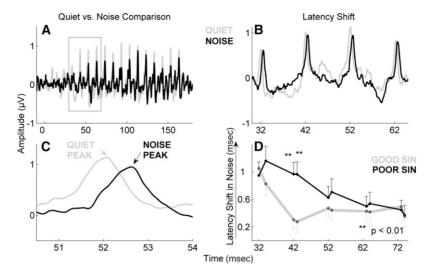
578

Figure 4. Effects of noise on brainstem responses in children with good and poor SIN perception. The effects are most evident in the transition region (*A,* boxed) of the response from 30 to 60 ms in the grand average waveforms of 66 children (*B* and *C*). Greater noise-induced latency shifts were noted in the children with poor SIN perception compared to children with good SIN perception ($p < 0.01$) *(D)*. Adapted from Anderson et al, 2010.

both poor SIN perceivers and children with dyslexia may be unable to benefit from stimulus predictabilities on a subcortical level, failing to make use of recent experience.

The Role of Brainstem Temporal Encoding in SIN Perception

Timing is an important feature in object identification and for perceptual discrimination. The differentiation of stop consonants is known to be especially challenging in the presence of background noise (Miller and Nicely, 1955). In order to evaluate the relationship between the subcortical representation of stop-consonant timing and SIN perception, cABRs to the syllables /ba/, /da/, and /ga/ were recorded in a group of children with a wide range of reading abilities (ages 8–14), including children with reading deficits (Hornickel et al, 2009). Children with reading disorders were included because of previous findings indicating that children with language-based learning disabilities have difficulty understanding speech in background noise (Bradlow et al, 2003; Ziegler et al, 2005). The auditory brainstem representation of formant frequencies that differ between these syllables was reflected in cABR timing differences, and the extent of these frequency differences correlated with SIN perception, with the best SIN perceivers having brainstem differentiation of the stop consonants that more closely follows the predicted pattern than that in the worst perceivers (Fig. 3).

The effects of background noise on ABRs are well-established and include delays in peak latencies and reductions in response amplitudes when compared to ABRs recorded in quiet conditions (Hall, 1992; Cunningham et al, 2001; Burkard and Sims, 2002). Such effects are particularly prevalent in the region of the response that corresponds to the formant transition in the speech syllable. In both children (Russo et al, 2004; Anderson et al, 2010) and older adults (Anderson et al, 2009), greater noise-induced shifts in peak latencies were found in poor SIN perceivers compared to good perceivers in quiet conditions when compared with responses recorded in background noise (Fig. 4). Thus, poor SIN perceivers are more vulnerable to noise-induced reductions in subcortical neural synchrony, likely leading to decreases in the temporal resolution that is required for accurate perception.

Musician and Linguistic Enhancement for SIN Perception

To better understand the interplay of sensory and cognitive functions in SIN perception, it is useful to examine the roles that both language and music play in the shaping of sensory activity by comparing typical and expert populations (e.g., musicians, tonal language speakers). For example, adult native speakers of Mandarin Chinese demonstrate more accurate representation of Mandarin rising and falling tones in the brainstem FFR compared to native speakers of English (Krishnan et al, 2009). Musicians have larger response amplitudes for encoding of both speech and music stimuli compared to nonmusicians (Musacchia et al, 2007) (Fig. 5). Similarly, musicians have more robust brainstem encoding of linguistically meaningful pitch contours compared to nonmusicians, indicating shared subcortical processing for speech and music as well as possible generalization of effects of corticofugal tuning from one domain to another (Wong et al, 2007). A musician advantage has been found for pitch, timing, and timbre representation
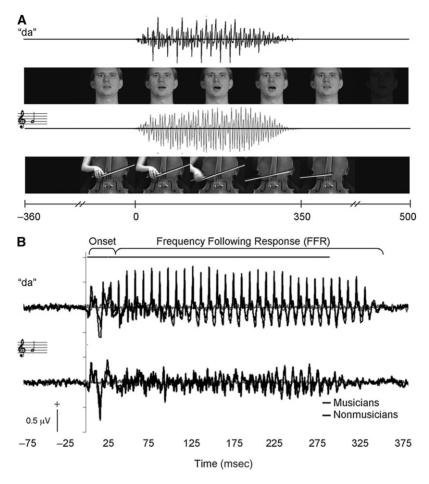
Figure 5. Stimulus timelines and audiovisual grand averages. *(A)* Auditory and visual components of speech and music stimuli. Acoustic onsets for both speech and music occurred 350 msec after the first video frame and simultaneously with the release of consonant closure and onset of string vibration, respectively. Speech and music sounds were 350 msec in duration and similar to each other in envelope and spectral characteristics. *(B)* Grand average brainstem responses to audiovisual speech *(upper)* and cello *(lower)* stimuli. Amplitude differences in the responses between musicians and controls are evident over the entire response waveforms ($p < 0.05$). Adapted from Musacchia et al, 2007.

in ABRs (Musacchia et al, 2007; Wong et al, 2007; Kraus et al, 2009; Lee et al, 2009; Strait et al, 2009a, 2009b). Moreover, the degree of subcortical enhancement varies with extent of musical experience, indicating that the musician advantage may stem, at least in part, from the modulating effects of life-long auditory experience rather than from innate neural characteristics.

Musical experience does not result in an overall gain effect but rather enhances the salient aspects of a signal. For example, in responses to musical chords, musicians have stronger responses for the higher harmonics and combination tones (important for melody recognition) but not for the fundamental frequency (Lee et al, 2009). This selective enhancement is also seen in the encoding of vocal emotion in a baby's cry (Strait et al, 2009b), with musicians showing greater processing efficiency through enhanced representation of the most spectrally and temporally transient region of the stimulus, compared to the more periodic, acoustically stable region (Fig. 6).

The musician advantage extends to behavioral and subcortical processing of speech in noise (Parbery-Clark, Skoe, Kraus, 2009; Parbery-Clark, Skoe, Lam, et al, 2009). Musicians have years of experience attending to distinct streams of music in orchestras, bands, and other venues. This experience has led to improved auditory perceptual skills, such as pitch discrimination (Tervaniemi et al, 2005; Micheyl et al, 2006; Rammsayer and Altenmuller, 2006), and enhancement of N1 and P2 in cortical-evoked and magnetoencephalographic responses (Shahin et al, 2003; Kuriki et al, 2006). Parbery-Clark, Skoe, Lam, et al (2009) found that musicians had higher scores on the HINT and QuickSIN, due in part to enhanced auditory working memory abilities (composed of the Woodcock-Johnson III Numbers Reversed and Auditory Working Memory subtests [Woodcock et al, 2001]). Furthermore, a comparison of ABRs to the speech syllable /da/ in quiet to those recorded to /da/ in six-talker babble demonstrated greater noise-induced peak timing delays in nonmusicians
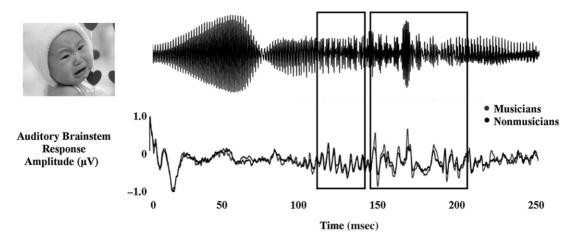
Figure 6. Stimulus (infant cry) and grand average response waveforms from musicians (gray) and nonmusicians (black). Response waveforms have been shifted back in time (7 msec) to align the stimulus and response onsets. Boxes delineate two stimulus subsections and the corresponding brainstem responses. The first subsection (112–142 msec) corresponds to the most periodic portion of the response and the corresponding region in the ABR. The second subsection (145–212 msec) corresponds to the more acoustically complex portion of the stimulus, characterized by transient amplitude bursts and rapid spectral changes. Musicians' responses demonstrate greater amplitudes than nonmusicians' responses throughout the complex region of the response (peak 1: $p < 0.003$; peak 2: $p < 0.03$) but not for the periodic region. Adapted from Strait et al, 2009a.

than in musicians, similar to the greater delays noted in children with poor SIN perception (Fig. 7).

Empirical study of musicians demonstrates the enhancement of sensory processing; moreover, this enhancement represents a selective rather than an overall gain effect (Chandrasekaran and Kraus, 2010a). The fact that musical experience enhances the ability to hear speech in challenging listening environments suggests that musical training may serve to enhance education in other domains, such as reading, and may provide an appropriate remediation strategy for individuals with impaired auditory processing.
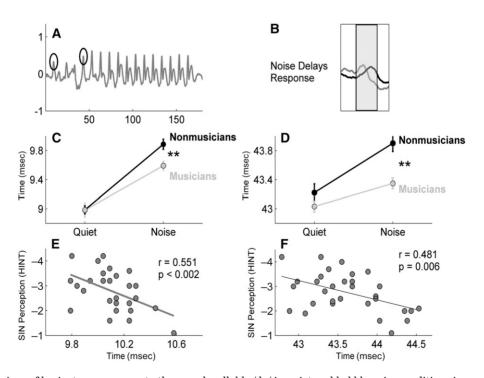

Figure 7. Comparison of brainstem responses to the speech syllable /da/ in quiet and babble noise conditions in musicians vs. nonmusicians. The selected peaks (onset and transition) are circled (A). Noise delays peak latencies (B), particularly in the onset and transition portions of the response. The musicians (gray) show significantly shorter latency delays in noise than nonmusicians (black) for the onset (C, $p < 0.01$) and transition peaks (D, $p < 0.01$). The latencies of the onset (E) and transition peaks (F) are correlated with SIN perception (onset: r = 0.551, $p < 0.002$; transition: r = 0.481, $p = 0.006$). Adapted from Parbery-Clark, Skoe, Kraus, 2009.

DISCUSSION

Successful communication in noisy environments involves speech processing at several stages. The sensory system, from the cochlea to the auditory cortex, must extract key features of the signal while suppressing irrelevant details. These features interact with cognitive processing, where sufficient working memory skills are needed to temporarily store this information while ignoring nonessential noise sources. The brainstem's particular roles include locking onto stimulus regularities to provide the cortex with a sharply tuned and stable representation of the stimulus. Other brainstem-level neural signatures important for successful SIN perception include robust encoding of the pitch and the preservation of temporal resolution in the presence of background noise. Cognitive and linguistic cues fill in the missing details.

Sensory-cognitive interactions are mediated by a massive corticofugal system (Suga and Ma, 2003). Brainstem responses to speech are shaped by both the acoustics of the incoming signal and cognitive processes such as attention and memory (Lukas, 1981; Bauer and Bayles, 1990; Galbraith et al, 1997; Galbraith et al, 1998). Auditory attention works to extract relevant signal elements from competing backgrounds and stores them in working memory (Johnson and Zatorre, 2005). These steps enable top-down predictive coding, thereby enhancing the brainstem encoding of relevant and/or predictable features (pitch, timing, and harmonics) (Ahissar and Hochstein, 2004; Kraus and Banai, 2007; Wong et al, 2007; de Boer and Thornton, 2008; Song, Skoe, et al, 2008; Chandrasekaran, Krishnan, et al, 2009). Enhanced subcortical function provides improved signal quality to the auditory cortex. Top-down sharpening of sensory fields has been noted in the cortex (Schreiner, 1998; Fritz et al, 2003; Fritz et al, 2005; Atiani et al, 2009), inferior colliculus (Gao and Suga, 2000), and the cochlear nucleus (Suga and Ma, 2003).

While peripheral deficits impair bottom-up encoding of stimulus features, attention and memory deficits impair the top-down predictive coding mechanism that tunes ABRs. These factors appear to intersect in a reciprocally interactive fashion. We are currently evaluating the interaction between peripheral, central, and cognitive factors in speech-in-noise perception in a group of older adults, including those with sensorineural hearing loss. Upon completion of this project we hope to have a better understanding of the roles contributed by these various factors.

CLINICAL IMPLICATIONS

The sensory-cognitive interactions involved in speech-in-noise processing emphasize the need to consider structures beyond the cochlea in evaluation and management of hearing difficulties. Behavioral measures used in the assessment of auditory processing disorders (usually manifested as difficulty with speech-in-noise understanding) can be affected by nonauditory variables, such as motivation, attention, and task difficulty. The cABR is an objective, noninvasive tool that provides information regarding the brainstem's ability to process the temporal and frequency features of the speech stimulus. Computer-based adaptive auditory training programs have been developed to facilitate learning through the use of exaggerated temporal cues and other strategies (Tallal, 2004; Sweetow and Sabes, 2006; Smith et al, 2009). Training-induced auditory brainstem plasticity has been documented (Russo et al, 2005; Song, Skoe, et al, 2008), and we are currently examining the effects of auditory training on brainstem encoding of speech in noise. The cABR may provide a clinically useful method for assessing the efficacy of auditory training as well as for identification of individuals who are most likely to benefit from auditory training or remediation. A clinical technology, BioMARK (Biologic Marker of Auditory Processing), is available as an addition to the Navigator Pro Auditory Evoked Potential hardware (Natus, Inc., San Carlos, CA). It was designed to quickly and objectively assess disorders of speech processing that may be present in children with language-based learning impairments, and normative data has been developed for children ages 3–4, 5–12, and 18–28 (Johnson et al, 2007; Song, Banai, et al, 2008; Banai et al, 2009; Dhar et al, 2009; Russo et al, 2009). The current BioMARK protocol requires approximately 20 min to implement, including time for electrode application and response analysis. It should be reasonable to use BioMARK to assess auditory function in individuals experiencing difficulty hearing in noise and to provide an objective metric of training-associated progress. Efforts are currently underway to establish normative data for infants as well as older adults with and without hearing loss.

The role of lifelong experience in shaping behavioral and neural measures of SIN perception indicates the need to take into account a broader range of life factors in patient case histories, particularly focusing on the history of musical training and/or language learning. Because speech and music share neural processing pathways and involve a myriad of common sensory and cognitive functions, the inclusion of musical components into auditory training programs may enhance motivation as well as functional outcomes.

REFERENCES

Ahissar M. (2007) Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn Sci* 11:458–465.

Ahissar M, Hochstein S. (2004) The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci* 8:457–464.

Ahissar M, Lubin Y, Putter-Katz H, Banai K. (2006) Dyslexia and the failure to form a perceptual anchor. *Nat Neurosci* 9:1558–1564.

Aiken SJ, Picton TW. (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245:35–47.

Anderson S, Parbery-Clark A, Skoe E, Kraus N. (2009) *Brainstem correlates of speech in noise in older adults.* Poster presented at Aging and Speech Communication: Third International and Interdisciplinary Research Conference, Indiana University, Bloomington, IN, October 12–14.

Anderson S, Skoe E, Chandrasekaran B, Kraus N. (2010) Neural timing is linked to speech perception in noise. *J Neurosci* 30:4922–4926.

Anderson S, Skoe E, Chandrasekaran B, Zecker S, Kraus N. (2010) Brainstem correlates of speech-in-noise perception in children. *Hear Res* 270:151–157.

Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA. (2009) Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61:467–480.

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. (2009) Reading and subcortical auditory function. *Cereb Cortex* 19:2699–2707.

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

Bauer LO, Bayles RL. (1990) Precortical filtering and selective attention: an evoked potential analysis. *Biol Psychol* 30:21–33.

Bee MA, Klump GM. (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. *J Neurophysiol* 92:1088–1104.

Best V, Gallun FJ, Carlile S, Shinn-Cunningham BG. (2007) Binaural interference and auditory grouping. *J Acoust Soc Am* 121:1070–1076.

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.

Bregman AS. (1990) *Auditory Scene Analysis.* Cambridge: MIT.

Bregman A, McAdams S. (1994) Auditory scene analysis: the perceptual organization of sound. *J Acoust Soc Am* 95:1177–1178.

Brokx JP, Nooteboom S. (1982) Intonation and the perceptual separation of simultaneous voices. *J Phon* 10:23–26.

Bronkhorst A. (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *ACUSTICA Acta Acustica* 86:117–128.

Burkard RF, Sims D. (2002) The human auditory brain-stem response to high click rates: aging effects. *Am J Audiol* 11:13–22.

Caspary DM, Schatteman TA, Hughes LF. (2005) Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: role of inhibitory inputs. *J Neurosci* 25:10952–10959.

Chandrasekaran B, Hornickel J, Skoe E, Nicol TG, 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. (2010a) Music, noise-exclusion, and learning disabilities. *Music Percept* 27:297–306.

Chandrasekaran B, Kraus N. (2010b) The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47:236–246.

Chandrasekaran B, Krishnan A, Gandour JT. (2009) Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain Lang* 108:1–9.

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.

Darwin CJ, Hukin RW. (2000) Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J Acoust Soc Am* 107:970–977.

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.

Dhar S, Abel R, Hornickel J, et al. (2009) Exploring the relationship between physiological measures of cochlear and brainstem function. *Clin Neurophysiol* 120:959–966.

Frisina DR, Frisina RD. (1997) Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res* 106:95–104.

Fritz J, Elhilali M, Shamma S. (2005) Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hear Res* 206:159–176.

Fritz J, Shamma S, Elhilali M, Klein D. (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223.

Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM. (1995) Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport* 6:2363–2367.

Galbraith GC, Bhuta SM, Choate AK, Kitahara JM, Mullen TAJ. (1998) Brain stem frequency-following response to dichotic vowels during attention. *Neuroreport* 9:1889–1893.

Galbraith GC, Jhaveri SP, Kuo J. (1997) Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. *Electroencephalogr Clin Neurophysiol* 102:46–53.

Gao E, Suga N. (2000) Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc Natl Acad Sci U S A* 97:8081–8086.

Hall J. (1992) *Handbook of Auditory Evoked Responses.* Needham Heights, MA: Allyn and Bacon.

Heinrich A, Schneider BA, Craik FI. (2007) Investigating the influence of continuous babble on auditory short-term memory performance. *Q J Exp Psychol* 61:735–751.

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.

Humes LE. (1996) Speech understanding in the elderly. *J Am Acad Audiol* 7:161–167.

Jeng F-C, Hu J, Dickman B, Lin C-Y, Lin C-D, Wang C-Y, et al. (2010) Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. *Int J Audiol* 50(1):14–26.

Johnson JA, Zatorre RJ. (2005) Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cereb Cortex* 15:1609–1620.

Johnson K, Nicol TG, Zecker SG, Kraus N. (2007) Auditory brainstem correlates of perceptual timing deficits. *J Cogn Neurosci* 19:376–385.

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. *J Acoust Soc Am* 116:2395–2405.

Kim S, Frisina RD, Mapes FM, Hickman ED, Frisina DR. (2006) Effect of age on binaural speech intelligibility in normal hearing adults. *Speech Commun* 48:591–597.

Kraus N, Chandrasekaran B. (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605.

Kraus N, Banai K. (2007) Auditory-processing malleability: focus on language and music. *Curr Dir Psychol Sci* 16:105–110.

Kraus N, Nicol T. (2005) How can the neural encoding and perception of speech be improved? In: Merzenich M, Syka J, eds. *Plasticity and Signal Representation in the Auditory System*. New York: Spring, 259–270.

Kraus N, Skoe E, Parbery-Clark A, Ashley R. (2009) Experience-induced malleability in neural encoding of pitch, timbre, and timing: implications for language and music. *Ann N Y Acad Sci* 1169:543–557.

Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. (2009) Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport* 20:408–413.

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

Kuriki S, Kanda S, Hirata Y. (2006) Effects of musical experience on different components of meg responses elicited by sequential piano-tones and chords. *J Neurosci* 26:4046–4053.

Lee KM, Skoe E, Kraus N, Ashley R. (2009) Selective subcortical enhancement of musical intervals in musicians. *J Neurosci* 29:5832–5840.

Lukas JH. (1981) The role of efferent inhibition in human auditory attention: an examination of the auditory brainstem potentials. *Int J Neurosci* 12:137–145.

Meddis R, O'Mard L. (1997) A unitary model of pitch perception. *J Acoust Soc Am* 102:1811–1820.

Micheyl C, Carlyon RP, Gutschalk A, et al. (2007) The role of auditory cortex in the formation of auditory streams. *Hear Res* 229:116–131.

Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. (2006) Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res* 219:36–47.

Miller GA, Nicely PE. (1955) An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27:338–352.

Moore BC, Peters RW, Glasberg BR. (1985) Thresholds for the detection of inharmonicity in complex tones. *J Acoust Soc Am* 77:1861–1867.

Musacchia G, Sams M, Skoe E, Kraus N. (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci U S A* 104:15894–15898.

Oxenham AJ. (2008) Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif* 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. *J Neurosci* 29:14100–14107.

Parbery-Clark A, Skoe E, Lam C, Kraus N. (2009) Musician enhancement for speech-in-noise. *Ear Hear* 30:653–661.

Parikh G, Loizou PC. (2005) The influence of noise on vowel and consonant cues. *J Acoust Soc Am* 118:3874–3888.

Rammsayer T, Altenmuller E. (2006) Temporal information processing in musicians and nonmusicians. *Music Percept* 24:37–48.

Russo N, Nicol T, Musacchia G, Kraus N. (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 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. *Dev Sci* 12:557–567.

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.

Sayles M, Winter IM. (2008) Ambiguous pitch and the temporal representation of inharmonic iterated rippled noise in the ventral cochlear nucleus. *J Neurosci* 28:11925–11938.

Schneider BA, Pichora-Fuller MK. (2001) Age-related changes in temporal processing: implications for speech perception. *Semin Hear* 22:227–240.

Schreiner CE. (1998) Spatial distribution of responses to simple and complex sounds in the primary auditory cortex. *Audiol Neuro-otol* 3:104–122.

Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. (2003) Enhancement of neuroplastic p2 and n1c auditory evoked potentials in musicians. *J Neurosci* 23:5545–5552.

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* 31:302–324.

Smith GE, Housen P, Yaffe K, et al. (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.

Snyder JS, Alain C. (2007) Toward a neurophysiological theory of auditory stream segregation. *Psychol Bull* 133:780–799.

Song JH, Banai K, Kraus N. (2008) Brainstem timing deficits in children with learning impairment may result from corticofugal origins. *Audiol Neurootol* 13:335–344.

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, Nicol TN, Kraus N. (In press) Test-retest reliability of the speech-evoked auditory brainstem response. *Clin Neurophysiol*.

Song JH, Skoe E, Banai K, Kraus N. (In press) Perception of speech in noise: neural correlates. *J Cogn Neurosci*.

Song JH, Skoe E, Wong PCM, Kraus N. (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20:1892–1902.

Sperling AJ, Zhong-Lin L, Manis FR, Seidenberg MS. (2005) Deficits in perceptual noise exclusion in developmental dyslexia. *Nat Neurosci* 8:862–863.

Strait D, Kraus N, Skoe E, Ashley R. (2009a) Musical experience promotes subcortical efficiency in processing emotional vocal sounds. *Ann N Y Acad Sci* 1169:209–213.

Strait DL, Kraus N, Skoe E, Ashley R. (2009b) Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *Eur J Neurosci* 29: 661–668.

Suga N, Ma X. (2003) Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat Rev Neurosci* 4:783–794.

Sweetow RW, Sabes JH. (2006) The need for and development of an adaptive Listening and Communication Enhancement (lace) program. *J Am Acad Audiol* 17:538–558.

Tallal P. (2004) Improving language and literacy is a matter of time. *Nat Rev Neurosci* 5:721–728.

Tervaniemi M, Just V, Koelsch S, Widmann A, Schröger E. (2005) Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp Brain Res* 161:1–10.

Tremblay K, Piskosz M, Souza P. (2003) Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol* 114:1332–1343.

Tzounopoulos T, Kraus N. (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62: 463–469.

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. *Biol Psychol* 67: 299–317.

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

Woodcock RW, McGrew KS, Mather N. (2001) *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Ziegler JC, Pech-Georgel C, George F, Alario F-X, 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.