

A Neural Basis of Speech-in-Noise Perception in Older Adults

Samira Anderson,^{1,2} Alexandra Parbery-Clark,^{1,2} Han-Gyol Yi,^{1,2} and Nina Kraus¹⁻⁴

Objective: We investigated a neural basis of speech-in-noise perception in older adults. Hearing loss, the third most common chronic condition in older adults, is most often manifested by difficulty understanding speech in background noise. This trouble with understanding speech in noise, which occurs even in individuals who have normal-hearing thresholds, may arise, in part, from age-related declines in central auditory processing of the temporal and spectral components of speech. We hypothesized that older adults with poorer speech-in-noise (SIN) perception demonstrate impairments in the subcortical representation of speech.

Design: In all participants (28 adults, age 60–73 yr), average hearing thresholds calculated from 500 to 4000 Hz were ≤ 25 dB HL. The participants were evaluated behaviorally with the Hearing in Noise Test (HINT) and neurophysiologically using speech-evoked auditory brainstem responses recorded in quiet and in background noise. The participants were divided based on their HINT scores into top and bottom performing groups that were matched for audiometric thresholds and intelligent quotient. We compared brainstem responses in the two groups, specifically, the average spectral magnitudes of the neural response and the degree to which background noise affected response morphology.

Results: In the quiet condition, the bottom SIN group had reduced neural representation of the fundamental frequency of the speech stimulus and an overall reduction in response magnitude. In the noise condition, the bottom SIN group demonstrated greater disruption in noise, reflecting reduction in neural synchrony. The role of brainstem timing is particularly evident in the strong relationship between SIN perception and quiet-to-noise response correlations. All physiologic measures correlated with SIN perception.

Conclusion: Adults in the bottom SIN group differed from the audiometrically matched top SIN group in how speech was neurally encoded. The strength of subcortical encoding of the fundamental frequency appears to be a factor in successful speech-in-noise perception in older adults. Given the limitations of amplification, our results suggest the need for inclusion of auditory training to strengthen central auditory processing in older adults with SIN perception difficulties.

(*Ear & Hearing* 2011;32;750–757)

INTRODUCTION

Older adults* frequently report having difficulty understanding speech in background noise. Peripheral hearing loss explains some of this difficulty; in fact, hearing impairment ranks third among chronic conditions in older adults (aged 65 yr or older) after hypertension and arthritis (Yueh et al. 2003). Yet, speech-in-noise (SIN) perception difficulty can be present in older adults who have normal audiometric thresholds (Gordon-Salant & Fitzgibbons 1993; Cruickshanks et al. 1998). The audiogram fails to accurately predict speech recognition per-

formance for all age groups, particularly in background noise (Killion & Niquette 2000; Souza et al. 2007). The relationship between pure-tone thresholds and SIN perception is even weaker in the older population (Dubno et al. 1984; Hargus & Gordon-Salant 1995; Kim et al. 2006), with word recognition scores in background noise falling below predicted scores by increasing amounts from middle to older adult ages (Souza et al. 2007).

The factors contributing to SIN perception difficulties are not well understood. Deficits in cognitive processing contribute to the problems experienced by older adults when listening in noise (Frisina & Frisina 1997; Gordon-Salant & Fitzgibbons 1997; Tremblay et al. 2002; Tun et al. 2002). For example, older adults are more vulnerable to the distracting effects of semantic content in background noise than younger adults, indicating that cognitive factors contribute to their SIN perception difficulties (Tun et al. 2002). Age-related sensory declines, including those in SIN perception, may be accompanied by increased activation in more general cognitive regions (i.e., working memory and attention) as a means of compensating for these declines (Wong et al. 2009). Therefore, older adults who have experienced declines in memory or attention are especially affected by decreases in sensory perception (Shinn-Cunningham & Best 2008).

SIN perception may also be affected by changes in central auditory processing. Aging affects the ability to process pitch cues (Helfer & Vargo 2009). Given that pitch is an important cue for speaker identification and object formation, this deficit may interfere with the older listener's ability to follow a single voice from among a stream of competing voices (Oxenham 2008; Shinn-Cunningham & Best 2008). Older adults are unable to benefit from voicing cues as effectively as younger adults in an informational masking task (Helfer & Freyman 2008; Huang et al. 2010). A possible deficit in frequency representation is supported by reduced frequency following responses (FFRs) to tone bursts, as well as increased frequency discrimination difference limens in older adults compared with younger adults (Clinard et al. 2010).

Age-related changes have been found in temporal processing. Declines in temporal resolution, as measured by gap detection, have been found in both animal (Walton et al. 1998) and human studies on aging (He et al. 1999; Harris et al. 2010; Ross et al. 2010; Walton 2010), and gap detection thresholds are related to SIN perception (Gordon-Salant & Fitzgibbons 1993; Feng et al. 2010). Precise representation of neural timing is important for capturing fast changing acoustic transitions that characterize consonants. Older adults demonstrate auditory brainstem response offset timing delays to speech syllables compared with younger adults (Vander Werff & Burns 2011). Furthermore, older adults have impaired neural representation of voice-onset-time contrasts, indicating that some of the SIN perception difficulties experienced by older adults may arise

¹Auditory Neuroscience Laboratory, and Departments of ²Communication Sciences, ³Neurobiology and Physiology, and ⁴Otolaryngology, Northwestern University, Evanston, Illinois.

*Defined as age 60 yr or older for the purposes of this study.

TABLE 1. The means and SDs for top and bottom SIN groups as well as *p* values for group differences are listed for HINT-Front scores, pure-tone averages (0.5–4 kHz), age, click wave V latencies, WASI IQ standard scores, and Woodcock-Johnson III Working Memory standard scores

	HINT dB (SNR)	PTA for 0.5–4 kHz (dB HL)	Age (yr)	Click Latency Wave V (msec)	WASI IQ (Standard Score)	Working Memory (Standard Score)
Total Group, mean (SD)	–2.69 (0.92)	13.71 (6.01)	63 (3.83)	6.04 (.31)	122 (11.21)	115 (14.74)
Top SIN, mean (SD)	–3.40 (0.54)	14.43 (2.37)	62 (3.24)	6.03 (.33)	123 (11.32)	111 (13.64)
Bottom SIN, mean (SD)	–1.99 (0.64)	12.99 (4.21)	64 (4.31)	6.04 (.31)	120 (10.67)	119 (14.97)
<i>P</i>	<0.001	0.538	0.265	0.959	0.391	0.119

from impaired temporal resolution in the central auditory system (Tremblay et al. 2003). Accurate subcortical representation of temporal information (as assessed by peak timing) is known to contribute to SIN perception in young adults and children (Kraus et al. 1995; Hornickel et al. 2009; Parbery-Clark et al. 2009a; Tzounopoulos & Kraus 2009; Anderson et al. 2010a; Hornickel et al. 2011; Skoe et al. 2011). Therefore, age-related decreases in temporal processing may lead to impairments in SIN perception.

The auditory brainstem response to complex sounds (cABR) is ideal for the assessment of auditory processing mechanisms associated with SIN perception (as reviewed in Anderson & Kraus 2010). The clear relationship between the stimulus and brainstem response waveforms allows for the direct comparison of the frequency and timing components of the stimulus and the response (Galbraith et al. 1995; Skoe & Kraus 2010). Previous work has demonstrated relationships between SIN perception and the temporal and spectral components of the cABR in both children (Chandrasekaran et al. 2009; Hornickel et al. 2009; Anderson et al. 2010a, b; Hornickel et al. 2011; Skoe et al. 2011) and young adults (Parbery-Clark et al. 2009a; Song et al. 2010). Subcortical encoding of the fundamental frequency (F_0) of the stimulus, an important cue in the perception of pitch (Fellows et al. 1997; Meddis & O'Mard 1997), appears to play a role in SIN perception, with better SIN perception linked to greater F_0 magnitudes in the cABR (Anderson et al. 2010b; Song et al. 2010). Likewise, greater degradation of brainstem responses to speech presented in background noise, as evidenced by delayed peak timing and reduced morphology, is associated with poor SIN perception (Parbery-Clark et al. 2009a; Anderson et al. 2010a, b). We therefore hypothesized that older adults with better SIN perception would have greater representation of the F_0 in the cABR and would have less overall degradation of neural timing in the presence of noise than an audiometrically matched group with poorer SIN perception.

PARTICIPANTS AND METHODS

Participants

Twenty-eight older participants (age 60–73 yr; 15 females; mean = 63.1; SD = 3.83) were recruited from Chicago, IL, and surrounding communities. Audiometric air conduction thresholds were obtained at octave intervals from 125 to 12,000 Hz and bone conduction thresholds were obtained from 500 to 4000 Hz. In all participants, pure-tone averages (calculated as the average threshold from 500 to 4000 Hz) were ≤ 25 dB HL. No individual threshold was >40 dB at or below 4000 Hz or >60 dB above 4000 Hz, and conductive hearing loss (air-bone

gap ≥ 15 dB) was not present at two or more frequencies in either ear. Participants had no history of neurological disorders nor asymmetric pure-tone thresholds (defined as >15 dB difference at two or more frequencies between ears) and had normal click-evoked auditory brainstem responses (defined as a wave V latency of ≤ 6.8 msec at 80 dB SPL presented at a rate of 31.25 Hz). All participants had normal cognitive abilities as measured by intelligent quotients ≥ 95 on the Wechsler Abbreviated Scale of Intelligence (WASI; Zhu & Garcia 1999). All procedures were approved by the Northwestern University Institutional Review Board.

Speech in Noise

SIN perception was assessed using the Hearing in Noise Test (HINT; Bio-logic Systems Corp., Mundelein, IL). HINT is an adaptive, widely used clinical test that uses phonetically balanced Bamford-Kowal-Bench (BKB) sentences (Bench et al. 1979) superimposed on a fixed speech-shaped noise masker. The sentences are presented at varying signal-to-noise ratios (SNRs). The speech reception threshold is determined by the SNR at which the participants repeat the sentences correctly 50% of the time. Lower and more negative SNR thresholds indicate the ability to repeat sentences in higher levels of noise. Both the sentences and speech-shaped noise are delivered binaurally through insert earphones. Participants with hearing loss (thresholds >20 dB HL at any frequency from 250 to 8000 Hz; $N = 12$) were tested in the standard unaided condition as well as the aided condition supplied by the HINT software. The aided condition compensated for reduced audibility by applying gain to the sentences (National Acoustics Laboratories-Revised; Byrne & Dillon 1986) according to each individual's air conduction thresholds.

Participant Groups

The participants were grouped according to their HINT scores using aided scores when applicable for those with hearing loss. The participants' HINT scores formed a normal distribution around the mean of -2.69 dB SNR (SD = 0.92). Fourteen participants (7 females) who scored better than the median (< -2.75 dB SNR) were placed in the top SIN group and the other 14 (8 females) who performed worse (> -2.75 dB SNR) in the bottom SIN group. The two groups did not differ in age, working memory (cluster score obtained from two subtests of the Woodcock-Johnson III Test of Cognitive Abilities, Numbers Reversed and Auditory Working Memory; Woodcock et al. 2001), intelligent quotient, pure-tone average, click-ABR wave V latency, or hearing thresholds (see Table 1 for means and SDs and Fig. 1 for audiogram displaying mean audiometric thresholds for each group).

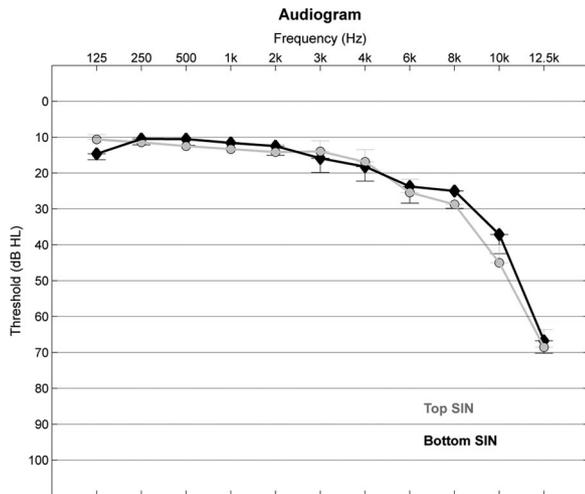


Fig. 1. Mean pure-tone thresholds (average of right and left) for top and bottom SIN groups from 125 to 12.5 kHz. No significant group differences were noted at any frequency. Error bars equal 1 standard error of measurement (SEM).

Stimuli for Neurophysiologic Recordings

For the cABR procedure, the speech stimulus was a syllable [da] synthesized at 20 kHz using a Klatt synthesizer (Klatt 1980) with a duration of 170 msec. This syllable consisted of a 50 msec transition (from the stop burst of [d] to [a]) followed by a 120 msec unchanging steady state region corresponding to the vowel [a]. After an initial 5 msec stop burst, the voicing remained constant at the F_0 of 100 Hz. During the consonant-vowel transition, the lower three formants shifted linearly (F_1 : 400–720 Hz; F_2 : 1700–1240 Hz; F_3 : 2580–2500 Hz) and then reached a plateau during the steady state region, whereas the upper three remained constant (F_4 : 3330 Hz; F_5 : 3750 Hz; F_6 : 4900 Hz) for the entire syllable.

The background noise stimulus was created from syntactically correct, nonsense English sentences spoken by six talkers (four females) in a conversational style. This 45 msec track (16 kHz sampling rate) was looped for the duration of the condition without pause. Consequently, the onset of the [da] coincided with different regions of the babble over the course of the condition. The [da] (with an interstimulus interval of 83 msec) was played at a SNR of +10 dB relative to the 70 dB SPL noise. Both the [da] and the noise stimuli were presented to the participants binaurally through insert earphones (ER-3, Etymotic Research, Elk Grove Village, IL) using NeuroScan Stim2 stimulus presentation software (NeuroScan Compumedics Inc., Charlotte, NC). See Figure 2 for the stimulus waveform and grand average responses and spectra to the stimulus presented in quiet and in noise conditions. To compensate for an individual's hearing loss (defined as thresholds >20 dB at any frequency from 250 to 6000 Hz), the NAL-R formula was used to selectively amplify frequencies of the stimulus based on the individual's hearing thresholds. Stimulus compensation was performed by MATLAB-coded routines (The MathWorks, Inc., Natick, MA) to create binaural stimuli amplified individually for each ear. Brainstem responses were collected with a binaural vertical montage with the active electrode placed at Cz, linked earlobes serving as reference, and forehead as the ground electrode. Responses were recorded using NeuroScan Acquire 4 at a sampling rate of 20 kHz. To maintain participant

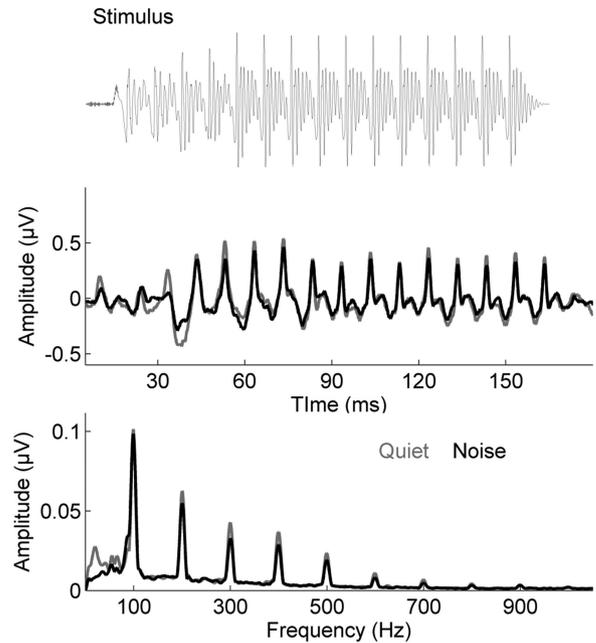


Fig. 2. Top: The stimulus waveform of the speech syllable [da]. Middle: Grand average response waveforms to the speech syllable [da] (N = 28) in quiet (gray) and noise (black). Bottom: Grand average frequency spectra of brainstem responses to the speech syllable [da] in quiet (gray) and noise (black) contain energy at the fundamental frequency (F_0 = 100 Hz) and integer multiples up to 1000 Hz.

cooperation and a relaxed state, participants watched muted subtitled movies of their choice. The recording sessions lasted for approximately 2 hr. All brainstem responses were collected in a sound-proof, electrically shielded booth.

The [da] stimulus was presented in alternating, opposite polarities. After obtaining subaverage responses to each polarity, the subaverages were added together to minimize contamination of the data by the cochlear microphonic and stimulus artifact (Gorga et al. 1985; Russo et al. 2004). This adding method, aside from minimizing noise, also strengthens the envelope following response (Aiken & Picton 2008). However, the response to the high-frequency content of the stimulus may be underrepresented in the added response (Aiken & Picton 2008). Therefore, we also subtracted the response to the two polarities to enhance spectral components that are canceled when the polarities are added (Aiken & Picton 2008; Heinz & Swaminathan 2009; Skoe & Kraus 2010). All analyses were conducted using added polarities unless indicated.

Data Analysis

Because the electrophysiological responses collected from Cz do not distinguish cortical and myogenic activities from those of subcortical nuclei, the EEG was filtered from 70 to 2000 Hz (12 dB/octave, zero phase-shift). This range reflects the upper but probably not the lower phase-locking function of the brainstem (Chandrasekaran & Kraus 2010). To analyze the evoked properties of the response, epoching was done from -40 to 213 msec relative to the stimulus onset at 0 msec. Any sweep exceeding $\pm 35 \mu\text{V}$ was considered an artifact and was excluded from the average. The initial 6000 accepted sweeps (3000 of each polarity) were included in the analysis for each listening condition.

Measurement of the Brainstem Response

All analyses were performed on the 5 to 190 msec window (the time range for the response to [da]). Average spectral amplitudes were calculated from the response, and zero padding was applied to the Fourier analysis to increase the resolution of the spectral display. For statistical analysis, the average amplitudes were calculated using 30 Hz bins around the frequencies of interest, which included the F_0 and its integer harmonics up to 1000 Hz. Root mean square (RMS) amplitude was used to quantify the overall magnitude of the response. The SNR of the response was measured by dividing the RMS of the response region (5–190 msec) of the waveform by the RMS of the prestimulus region (–40 to 0 msec).

Using cross-correlation, the similarity between the responses from the two conditions (quiet and noise) was assessed. Correlation coefficients were calculated by shifting the response waveform obtained in noise relative to the response waveform obtained in quiet (± 2 msec) until a maximum correlation was achieved, resulting in a Pearson's r value, which we refer to as the quiet-to-noise response correlation value. When the correlation between the quiet and noise condition is high (approaching 1), noise is interpreted as having a minimal impact on the response, whereas lower correlations indicate that noise has a greater impact on the response (Parbery-Clark et al. 2009a). Fisher's transformation was used to convert the r values to z scores for statistical computations.

Statistical Analysis

All statistical analyses were conducted with SPSS software (SPSS Inc., Chicago, IL). We compared cABRs in quiet and noise conditions using a two-way mixed-model analysis of variance (ANOVA) with group (top SIN versus bottom SIN) serving as the between-group independent variable and condition (quiet versus noise) serving as the within-group independent variable. F_0 and RMS amplitudes served as dependent variables. We did not expect to see group differences in representation of stimulus harmonics and used a separate MANOVA to compare differences in harmonics ($H_2 - H_{10}$). MANOVAs were also used to assess group differences in the impact of noise on responses (quiet-to-noise response correlations) and overall noise levels (response SNRs). The Levene's Test for Equality of Variances revealed homogeneity of variance for HINT scores and all dependent variables. Pearson's correlations were used to explore the continuous relationships between the measures of SIN perception and brainstem encoding in quiet and noise.

RESULTS

Groups Differ in the F_0 and RMS Magnitudes

F_0 magnitudes were greater in the top SIN group than in the bottom SIN group ($F[1,26] = 6.595$; $p = 0.016$). The group difference was significant in quiet ($F[1,26] = 7.769$; $p = 0.010$; Fig. 3) and was trending toward significance in noise ($F[1,26] = 3.793$; $p = 0.062$; Fig. 4). Moreover, the top SIN group had greater RMS amplitudes compared with the bottom group ($F[1,26] = 5.003$; $p = 0.034$). The group differences in RMS were significant for both the quiet ($F[1,26] = 4.700$; $p = 0.039$; Fig. 3) and the noise conditions ($F[1,26] = 4.348$; $p =$

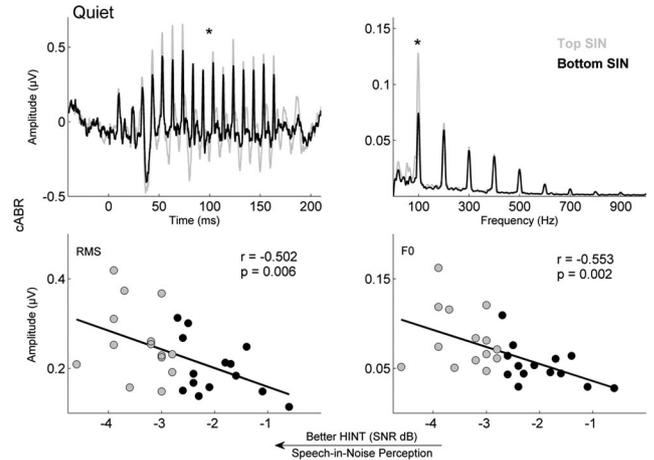


Fig. 3. Top left: Average responses for top (gray) and bottom (black) SIN groups to the speech syllable [da] in quiet. A significant RMS difference was noted ($*p = 0.039$). Top right: Average frequency spectra with significant difference for F_0 ($*p = 0.010$) calculated over the entire response range (5–190 msec). Bottom: HINT scores are significantly related to RMS (left) and F_0 (right) amplitudes.

0.047; Fig. 4). The overall interaction between group and condition was not significant ($F[2,25] = 0.883$; $p = 0.426$), indicating that noise had similar effects on both the top and bottom SIN groups for the F_0 and RMS measurements.

HINT Scores Are Related to F_0 and RMS Magnitudes

Better HINT scores (i.e., lower SNRs) were associated with higher F_0 magnitudes and RMS amplitudes. HINT scores were related to F_0 magnitude in quiet ($r = -0.553$, $p = 0.002$) and in noise ($r = -0.489$, $p = 0.008$) and RMS amplitudes ($r = -0.503$, $p = 0.006$) in quiet and in noise ($r = -0.528$, $p = 0.004$).

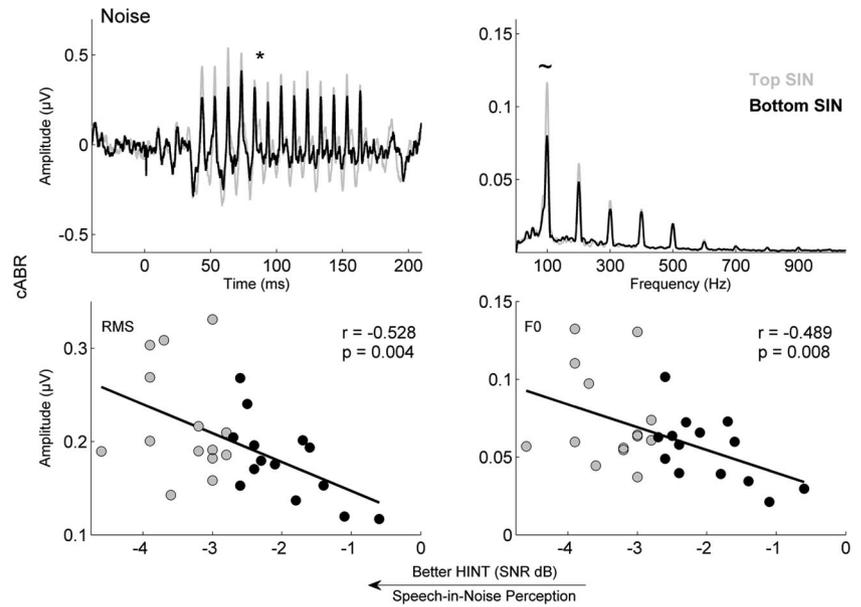
Groups Did Not Differ in the Prestimulus Range

The groups did not differ in the RMS of the prestimulus range (–40 to 0 msec) in quiet ($F[1,26] = 4.348$; $p = 0.047$) or in noise ($F[1,26] = 0.167$; $p = 0.687$), indicating that the differences in overall response amplitude were not influenced by prestimulus baseline neural activity. The groups also did not differ in response SNRs in quiet ($F[1,26] = 2.379$; $p = 0.135$) or in noise ($F[1,26] = 2.009$; $p = 0.168$).

Groups Did Not Differ in the Response to the Harmonics

There was no main effect of group for the response to the higher harmonics ($H_2 - H_{10}$) when they were entered as dependent variables in the MANOVA in quiet ($F[9,18] = 0.326$; $p = 0.955$) or in noise ($F[9,18] = 0.789$; $p = 0.630$). Subtracted polarities failed to reveal any group differences in quiet in the F_0 ($F[1,26] = 1.399$; $p = 0.248$) or higher harmonics ($H_2 - H_{10}$; $F[9,18] = 0.249$; $p = 0.981$), or in noise in the F_0 ($F[1,26] = 0.261$; $p = 0.614$) or higher harmonics ($H_2 - H_{10}$; $F[9,18] = 0.167$; $p = 0.995$). The groups differed in F_0 magnitude for added polarities but not for subtracted polarities, confirming that the differences arise from the brainstem response to the amplitude modulation of the signal.

Fig. 4. Top left: Average responses for top (gray) and bottom (black) SIN groups to the speech syllable [da] in six-talker babble. A significant RMS difference was noted ($*p = 0.047$). Top right: Average frequency spectra with the difference for F_0 approaching significance ($\sim p = 0.062$) when calculated over the entire range (5–190 msec). Bottom: HINT scores are significantly related to RMS (left) and F_0 (right) amplitudes.



Groups Differ in the Effects of Noise on Response Morphology and Timing

Quiet-to-noise response correlations revealed significant SIN group differences ($F[1,26] = 4.961$, $p = 0.035$; Fig. 5). The top SIN perceivers showed greater similarity between their responses in quiet and responses in noise than the bottom SIN perceivers. Correlation analyses indicated a strong relationship

between HINT scores and quiet-to-noise response correlation values ($r = -0.682$, $p < 0.001$; Fig. 5), in which better performance on the HINT was associated with reduced impact of noise on response morphology. To determine whether this relationship was driven by greater noise in the bottom SIN group's response, the correlation was repeated with the SNR in quiet entered as a covariate. The correlations remained strong even after controlling for the SNR in quiet ($r = -0.564$; $p = 0.002$).

DISCUSSION

Summary

Our results demonstrate the importance of subcortical encoding of the F_0 for successful SIN perception in older adults. Participants in the top SIN group had greater subcortical representation of F_0 magnitudes in response to a speech syllable than participants in the bottom SIN group. The top performing group exceeding the bottom group in RMS amplitudes. Furthermore, the neural response timing of participants with better SIN was less affected by noise, having higher quiet-to-noise correlations between responses. Strong relationships were found between subcortical response measures and behavioral measures of speech in noise (HINT scores).

Audiometrically Matched Groups Differ in SIN Perception

In our data set, two groups of older adults with matched audiograms differed in SIN performance, confirming previous work demonstrating that peripheral hearing does not fully account for the variance in SIN perception (Killion & Niquette 2000; Souza et al. 2007). People with hearing loss typically perform worse on perceptual SIN tests than those with normal hearing, yet some of our participants with moderate high-frequency hearing loss (above 3000 Hz) had speech-in-noise scores in the top 25th percentile of the overall group. We found that in these hearing-matched groups, brainstem encoding of the F_0 and the degree of noise-induced change on response morphology and timing were each predictive of SIN ability.

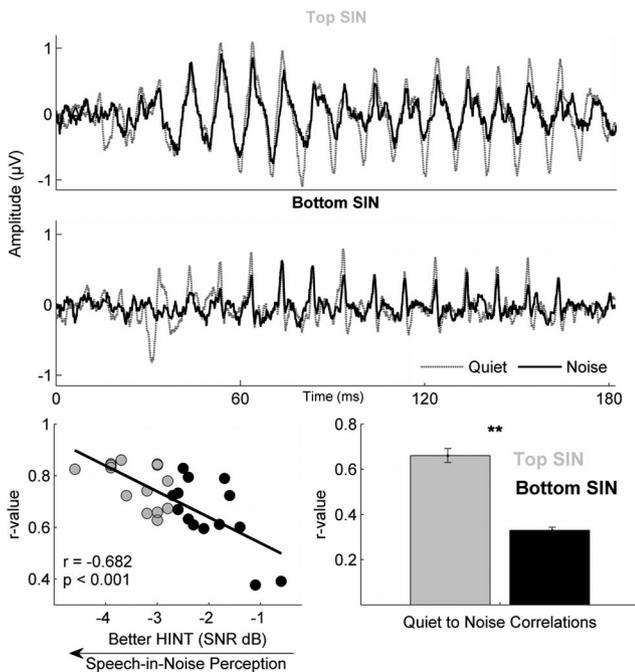


Fig. 5. Responses in the bottom SIN group changed significantly in noise compared with those in the top SIN group. Response waveforms in quiet (dashed lines) and noise (solid lines) are shown for an individual in the top SIN group (top panel) and the bottom SIN group (middle panel). Bottom panel, left: HINT scores are highly correlated with quiet-to-noise correlation r values ($p < 0.001$) over the entire response (5–190 msec). Bottom panel, right: Bar graphs illustrating significant group differences in quiet-to-noise correlation r values ($*p = 0.035$). Error bars equal 1 SEM.

These results highlight the need to consider subcortical encoding of speech as an important factor contributing to behavioral SIN performance.

Role of the F_0 Morphology and Timing in SIN Perception

In our data set, the subcortical encoding of the F_0 is an important factor in SIN perception. The F_0 and other pitch cues contribute to auditory object identification, allowing the listener to “tag” the target voice with a specific identity and to follow this particular voice from among competing voices or other noises. The ability to distinguish between competing streams of information is dependent, in part, on the F_0 , as demonstrated by enhanced discrimination of vowels with greater F_0 separation between concurrent vowels (Assmann & Summerfield 1987; Culling & Darwin 1993) and sentences (Brox & Nootboom 1982; Bird & Darwin 1998). Moreover, compared with younger adults, older adults benefit less from the pitch cues in male versus female contrasts (Helfer & Freyman 2008) or from prior familiarization with the target voice (Huang et al. 2010)—both key elements for establishing talker identity. Our results suggest that this lack of benefit may reflect subcortical deficits of pitch encoding. Age-related changes in perceptual measures involving processing of F_0 differences (Lam & Sanchez 2007; He et al. 2008; Clinard et al. 2010; Souza et al. 2011) and in a frequency discrimination task using cortical evoked potentials (Harris et al. 2008) support the idea of decreased neurophysiological representation of frequency in older adults.

Although pitch plays an important role in SIN perception, other aspects of the speech signal (e.g., timing and harmonics) are equally as important for understanding one voice from among a background of voices (Shinn-Cunningham & Best 2008). While we found no differences in the representation of individual harmonics among SIN perceivers, we do not interpret these results as indicating that harmonic encoding does not contribute to SIN perception. Rather, we conjecture that the representation of pitch and preservation of timing were important factors for distinguishing between good and poor SIN perceivers in this group of individuals. Furthermore, our results indicate that the neural representation of pitch does not fully account for SIN perception, suggesting that there are other important aspects of cortical and subcortical encoding which contribute to SIN perception in older adults. That is, response timing as reflected by quiet-to-noise correlation r values was indeed another factor. Harmonic information is likely to prove particularly important for individuals with hearing loss.

Age-Related Changes in GABA Inhibition

The decline in the ability to use pitch cues may arise from age-related decreases in γ -aminobutyric acid (GABA) inhibition. Decreases in GABA were found in the inferior colliculus and dorsal cochlear nucleus of rats (Caspary et al. 1995, 2005). Downregulation of inhibitory function may lead to degradation of subcortical temporal resolution (Caspary et al. 2008) by decreasing selectivity of pertinent acoustic features in the stimulus (Burger & Pollak 1998; Hall 1999; Edwards et al. 2008). We conjecture that GABA inhibition may be partly responsible for stronger F_0 encoding and more stable/precise neural timing in the top SIN group.

While decreases in GABAergic inhibition may contribute to age-related deficits in subcortical encoding of pitch and timing, the primary purpose of this study was to examine aspects of subcortical processing important for SIN perception in older adults rather than to assess the effects of aging on subcortical responses. Given previous findings demonstrating smaller representation of the F_0 in children and young adults with poor SIN perception (Anderson et al. 2010b; Song et al. 2010), it is possible that our replication of these effects in an older population is indicative of a fundamental mechanism of auditory processing—a process that would hence be age-independent. Whether similar mechanisms operate in the hearing impaired older adult is currently under investigation; the present work provides a foundation for considering the influence of this key factor on the neural representation of sound.

Clinical Implications

Peripheral aspects of presbycusis, such as widened auditory filters and reduced frequency selectivity (Florentine et al. 1980), and central factors, such as decreased neural synchrony and impaired temporal resolution, limit the benefits of amplification (Tremblay et al. 2003; Gordon-Salant 2005). Therefore, what are the implications of our findings for clinical management of people with hearing-in-noise difficulties? Given the limits of amplification, it is important to examine the effects of auditory training on SIN perception. Neural plasticity in the auditory pathway has been demonstrated in animal (Gao & Suga 2000; Bajo et al. 2010) and human studies (Tremblay et al. 2001; Tremblay & Kraus 2002; Russo et al. 2005; de Boer & Thornton 2008; Song et al. 2008). For example, training in pitch discrimination leads to more robust neural phase locking in the FFR (Song et al. 2008; Carcagno & Plack 2010). Auditory training also results in improved SIN perception based on behavioral and self-assessment measures (Sabes & Sweetow 2007; Burk & Humes 2008). Moreover, musicians, who undergo a form of life-long auditory training, have better SIN perception and enhanced subcortical encoding of speech in noise compared with nonmusicians (Parbery-Clark et al. 2009a, b; Bidelman & Krishnan 2010) and generally enhanced auditory skills (Kraus & Chandrasekaran 2010). In these studies, the impact of experience is evident in the very aspects of processing (neural timing and frequency representation, especially in noise) that distinguish our top and bottom SIN perceivers, demonstrating that auditory training (both short- and long-term) can be used to modulate responses in the auditory pathway. The determination of neural factors associated with better SIN perception may lead to more effective training programs.

Future Work

In our dataset, the personal characteristics of individuals with robust F_0 encoding were varied. At present, we cannot pinpoint the mechanisms or experiences contributing to enhanced brainstem responses in these older adults. Future work will use Structural Equation Modeling (a statistical technique for estimating causal relationships) to examine the contributions of physical fitness, and cognitive factors such as memory and attention, to the strength of subcortical responses and to performance on speech-in-noise measures. It will be important to extend this work to older adults with more severe hearing loss, for whom peripheral pathology complicates interpretation

of factors contributing to SIN perception. The present study provides a baseline from which we can investigate the effects of hearing loss on these processes and examine neural response changes elicited by auditory and cognitive training paradigms targeting improved auditory perception in older adults.

CONCLUSIONS

In summary, our findings suggest that central auditory processes contribute to the variance in SIN perception experienced by older adults. These results extend the work of previous studies demonstrating relationships between SIN ability and subcortical representation of speech in children and young adults (Parbery-Clark et al. 2009a; Anderson et al. 2010a, b; Song et al. 2010) to older adults who have clinically normal hearing. The responses of the top SIN group had more robust representation of the F_0 and response morphology, magnitude and timing and were less affected by noise. The brainstem response to speech is therefore an objective measure of neural responses to speech in noise and has the potential to improve assessment and management of SIN difficulties.

ACKNOWLEDGMENTS

The authors thank Trent Nicol, Jane Hornickel, and Dr. Judy Song for their helpful comments on the original manuscript and Erika Skoe for her critical review and comments on the revised manuscript. They also thank the people who participated in this study. The authors wish to acknowledge the work of Erika Skoe and Dr. Sumitrajit Dahr for their help in developing the MATLAB routine used for stimulus amplification.

This work was supported by the National Institutes of Health Grant R01-DC010016.

Address for correspondence: Samira Anderson, AuD, Auditory Neuroscience Laboratory, 2240 Campus Dr., Evanston, IL 60208, USA. E-mail: sba@u.northwestern.edu.

Received September 24, 2010; accepted April 24, 2011.

REFERENCES

- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hear Res*, *245*, 35–47.
- Anderson, S., & Kraus, N. (2010). Objective neural indices of speech-in-noise perception. *Trends Amplif*, *14*, 73–83.
- Anderson, S., Skoe, E., Chandrasekaran, B., et al. (2010a). Neural timing is linked to speech perception in noise. *J Neurosci*, *30*, 4922–4926.
- Anderson, S., Skoe, E., Chandrasekaran, B., et al. (2010b). Brainstem correlates of speech-in-noise perception in children. *Hear Res*, *270*, 151–157.
- Assmann, P. F., & Summerfield, Q. (1987). Perceptual segregation of concurrent vowels. *J Acoust Soc Am*, *82*, S120.
- Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci*, *13*, 253–260.
- Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Brit J Audiol*, *13*, 108–112.
- Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res*, *1355*, 112–125.
- Bird, J., & Darwin, C. J. (1998). Effects of a Difference in Fundamental Frequency in Separating Two Sentences. In A. R. Palmer, A. Rees, A. Q. Summerfield, et al. (Eds). *Psychophysical and Physiological Advances in Hearing* (pp. 263–269). London: Whurr.
- Brox, J. P., & Nootboom, S. (1982). Intonation and the perceptual separation of simultaneous voices. *J Phonetics*, *10*, 23–26.
- Burger, R. M., & Pollak, G. D. (1998). Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the inferior colliculus. *J Neurophysiol*, *80*, 1686–1701.
- Burk, M. H., & Humes, L. E. (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. *J Speech Lang Hear Res*, *51*, 759–771.
- Byrne, D., & Dillon, H. (1986). The national acoustic laboratories' new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear*, *7*, 257–265.
- Carcagno, S., & Plack, C. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol*, *12*, 89–100.
- Caspary, D. M., Ling, L., Turner, J. G., et al. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *J Exp Biol*, *211*, 1781–1791.
- Caspary, D. M., Milbrandt, J. C., Helfert, R. H. (1995). Central auditory aging: GABA changes in the inferior colliculus. *Exp Gerontol*, *30*, 349–360.
- Caspary, D. M., Schatteman, T. A., Hughes, L. F. (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., et al. (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.
- Clinard, C. G., Tremblay, K. L., Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hear Res*, *264*, 48–55.
- Cruikshanks, K. J., Wiley, T. L., Tweed, T. S., et al. (1998). Prevalence of hearing loss in older adults in beaver dam, wisconsin: The epidemiology of hearing loss study. *Am J Epidemiol*, *148*, 879–886.
- Culling, J. F., & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0 . *J Acoust Soc Am*, *93*, 3454–3467.
- 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. *J Neurosci*, *28*, 4929–4937.
- Dubno, J., Dirks, D., Morgan, D. (1984). Effects of age and mild hearing loss on speech recognition in noise. *J Acoust Soc Am*, *76*, 87–96.
- Edwards, C. J., Leary, C. J., Rose, G. J. (2008). Mechanisms of long-interval selectivity in midbrain auditory neurons: Roles of excitation, inhibition, and plasticity. *J Neurophysiol*, *100*, 3407–3416.
- Fellowes, J., Remez, R., Rubin, P. (1997). Perceiving the sex and identity of a talker without natural vocal timbre. *Percept Psychophys*, *59*, 839–840.
- Feng, Y., Yin, S., Kieffe, M., & Wang, J. (2010). Temporal resolution in regions of normal hearing and speech perception in noise for adults with sloping high-frequency hearing loss. *Ear Hear*, *31*, 115–125.
- Florentine, M., Buus, S., Scharf, B., et al. (1980). Frequency selectivity in normally-hearing and hearing-impaired observers. *J Speech Hear Res*, *23*, 646–669.
- Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hear Res*, *106*, 95–104.
- Galbraith, G. C., Arbagey, P. W., Branski, R., et al. (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. *Proc Natl Acad Sci USA*, *97*, 8081–8086.
- Gordon-Salant, S. (2005). Hearing loss and aging: New research findings and clinical implications. *J Rehab Research Dev*, *42*, 9–24.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear Res*, *36*, 1276.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *J Speech Lang Hear Res*, *40*, 423–431.

- Gorga, M., Abbas, P., & Worthington, D. (1985). Stimulus calibration in ABR measurements. In J. Jacobsen (Ed.), *The auditory brainstem response* (pp. 49–62). San Diego: College Hill Press.
- Hall, J. C. (1999). Gabaergic inhibition shapes frequency tuning and modifies response properties in the auditory midbrain of the leopard frog. *J Comp Physiol*, *185*, 479–491.
- Hargus, S. E., & Gordon-Salant, S. (1995). Accuracy of speech intelligibility index predictions for noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment. *J Speech Hear Res*, *38*, 234–243.
- Harris, K. C., Eckert, M. A., Ahlstrom, J. B., et al. (2010). Age-related differences in gap detection: Effects of task difficulty and cognitive ability. *Hear Res*, *264*, 21–29.
- Harris, K. C., Mills, J. H., He, N.-J., et al. (2008). Age-related differences in sensitivity to small changes in frequency assessed with cortical evoked potentials. *Hear Res*, *243*, 47–56.
- He, N.-J., Horwitz, A. R., Dubno, J. R., et al. (1999). Psychometric functions for gap detection in noise measured from young and aged subjects. *J Acoust Soc Am*, *106*, 966–978.
- He, N.-J., Mills, J. H., Ahlstrom, J. B., et al. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J Acoust Soc Am*, *124*, 3841–3849.
- Heinz, M., & Swaminathan, J. (2009). Quantifying envelope and fine-structure coding in auditory nerve responses to chimaeric speech. *J Assoc Res Otolaryngol*, *10*, 407–423.
- Helfer, K., & Freyman, R. (2008). Aging and speech-on-speech masking. *Ear Hear*, *29*, 87–98.
- Helfer, K. S., & Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *J Am Acad Audiol*, *20*, 264–271.
- Hornickel, J., Chandrasekaran, B., Zecker, S., et al. (2011). Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. *Behav Brain Res*, *216*, 597–605.
- Hornickel, J., Skoe, E., Nicol, T., et al. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci USA*, *106*, 13022–13027.
- Huang, Y., Xu, L., Wu, X., Li, L. (2010). The effect of voice cuing on releasing speech from informational masking disappears in older adults. *Ear Hear*, *31*, 579–583.
- Killion, M., & Niquette, P. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *Hear J*, *53*, 46–53.
- Kim, S., Frisina, R. D., Mapes, F. M., et al. (2006). Effect of age on binaural speech intelligibility in normal hearing adults. *Speech Commun*, *48*, 591–597.
- Klatt, D. (1980). Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am*, *67*, 971–995.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat Rev Neurosci*, *11*, 599–605.
- Kraus, N., McGee, T., Carrell, T. D., et al. (1995). Central auditory system plasticity associated with speech discrimination training. *J Cog Neurosci*, *7*, 25–32.
- Lam, E., & Sanchez, L. (2007). Evaluation of screening instruments for auditory processing disorder (APD) in a sample of referred children. *Aust N Z J Audiol*, *29*, 26–39.
- Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. *J Acoust Soc Am*, *102*, 1811–1820.
- Oxenham, A. J. (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. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci*, *29*, 14100–14107.
- Parbery-Clark, A., Skoe, E., Lam, C., et al. (2009b). Musician enhancement for speech-in-noise. *Ear Hear*, *30*, 653–661.
- Ross, B., Schneider, B., Snyder, J. S., et al. (2010). Biological markers of auditory gap detection in young, middle-aged, and older adults. *Plos ONE*, *5*, e10101.
- Russo, N., Nicol, T., Musacchia, G., et al. (2004). Brainstem responses to speech syllables. *Clin Neurophysiol*, *115*, 2021–2030.
- Russo, N., Nicol, T., Zecker, S., et al. (2005). Auditory training improves neural timing in the human brainstem. *Behav Brain Res*, *156*, 95–103.
- Sabes, J. H., & Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (lace (tm)) training. *Int J Audiol*, *46*, 374–383.
- Shinn-Cunningham, B. G., & 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.
- Skoe, E., Nicol, T., Kraus, N. (2011). Cross-phaseogram: Objective neural index of speech sound differentiation. *J Neurosci Methods*, *196*, 308–317.
- Song, J., Skoe, E., Banai, K., et al. (2010). Perception of speech in noise: Neural correlates. *J Cog Neurosci*, *2248*–2259.
- Song, J. H., Banai, K., Kraus, N. (2008). Brainstem timing deficits in children with learning impairment may result from corticofugal origins. *Audiol Neurotol*, *13*, 335–344.
- Souza, P., Arehart, K., Miller, C. W., et al. (2011). Effects of age on F0 discrimination and intonation perception in simulated electric and electroacoustic hearing. *Ear Hear*, *32*, 75–83.
- Souza, P., Boike, K., Witherell, K., et al. (2007). Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise. *J Am Acad Audiol*, *18*, 54–65.
- Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *J Speech Lang Hear Res*, *45*, 564–572.
- Tremblay, K., Kraus, N., McGee, T., et al. (2001). Central auditory plasticity: Changes in the N1–P2 complex after speech-sound training. *Ear Hear*, *22*, 79–90.
- Tremblay, K., Piskosz, M., Souza, P. (2002). Aging alters the neural representation of speech cues. *Neuroreport*, *13*, 1865–1870.
- 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.
- Tun, P., O'Kane, G., Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychol Aging*, *17*, 453–467.
- Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*, 463–469.
- Vander Werff, K. R., & Burns, K. S. (2011). Brain stem responses to speech in younger and older adults. *Ear Hear*, *32*, 168–180.
- Walton, J. P. (2010). Timing is everything: Temporal processing deficits in the aged auditory brainstem. *Hear Res*, *264*, 63–69.
- Walton, J. P., Frisina, R. D., & O'Neill, W. E. (1998). Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse. *J Neurosci*, *18*, 2764–2776.
- Wong, P. C. M., Jin, J. X., Gunasekera, G. M., et al. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, *47*, 693–703.
- Woodcock, R. W., McGrew, K. S., Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.
- Yueh, B., Shapiro, N., MacLean, C. H., et al. (2003). Screening and management of adult hearing loss in primary care: Scientific review. *J Am Med Assoc*, *289*, 1976–1985.
- Zhu, J., & Garcia, E. (1999). *The Wechsler Abbreviated Scale of Intelligence (WASI)*. New York: Psychological Corporation.