

# Effects of hearing loss on the subcortical representation of speech cues

Samira Anderson,<sup>a)</sup> Alexandra Parbery-Clark, Travis White-Schwoch, Sarah Drehobl, and Nina Kraus<sup>b)</sup>

Northwestern University, Auditory Neuroscience Laboratory, Communication Sciences, 2240 North Campus Drive, Evanston, Illinois 60208

(Received 19 September 2012; revised 10 March 2013; accepted 20 March 2013)

Individuals with sensorineural hearing loss often report frustration with speech being loud but not clear, especially in background noise. Despite advanced digital technology, hearing aid users may resort to removing their hearing aids in noisy environments due to the perception of excessive loudness. In an animal model, sensorineural hearing loss results in greater auditory nerve coding of the stimulus envelope, leading to a relative deficit of stimulus fine structure. Based on the hypothesis that brainstem encoding of the temporal envelope is greater in humans with sensorineural hearing loss, speech-evoked brainstem responses were recorded in normal hearing and hearing impaired age-matched groups of older adults. In the hearing impaired group, there was a disruption in the balance of envelope-to-fine structure representation compared to that of the normal hearing group. This imbalance may underlie the difficulty experienced by individuals with sensorineural hearing loss when trying to understand speech in background noise. This finding advances the understanding of the effects of sensorineural hearing loss on central auditory processing of speech in humans. Moreover, this finding has clinical potential for developing new amplification or implantation technologies, and in developing new training regimens to address this relative deficit of fine structure representation.

© 2013 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4799804>]

PACS number(s): 43.71.Qr, 43.64.Qh, 43.71.Ky, 43.71.Lz [ELP]

Pages: 3030–3038

## I. INTRODUCTION

Individuals with sensorineural hearing loss (SNHL) often report that increases in speech intensity do not improve speech intelligibility, especially in background noise, suggesting that the nature of the communication problem extends beyond reduced audibility. A number of mechanisms may be responsible for the perception of loudness without clarity that accompanies SNHL. Most cases of SNHL involve the loss of cochlear outer hair cells, also known as the cochlear amplifier (Dallos *et al.*, 2006), that regulate the amplification of transmitted sounds via afferent and efferent connections (Lopez-Poveda *et al.*, 2005). The loss of outer hair cells results in a change from nonlinear compression, as is found in functioning outer hair cells (Robles *et al.*, 1986), to linear processing, resulting in steeper growth of loudness (Jürgens *et al.*, 2011).

The consequences of SNHL are not solely peripheral, however; deprivation of auditory input may lead to tonotopic remapping of the inferior colliculus (Wyllott, 1991) and other changes in central auditory processing (Aizawa and Eggermont, 2006), including a disruption in the gain mechanism in the auditory cortex (Morita *et al.*, 2003; Wienbruch *et al.*, 2006) and decreased auditory acoustic reflex thresholds (Munro and Blount, 2009). This disruption may arise

from alterations in the balance of excitatory and inhibitory neurotransmitters in the auditory brainstem (Dong *et al.*, 2010) and cortex (Kotak *et al.*, 2005), as demonstrated by induced SNHL in animal models and subsequent global increases in neural firing. In everyday situations, the proposed effects of these disruptions is that individuals with SNHL perceive loud sounds as loud as—or even louder than—someone with normal hearing (NH) (Moore *et al.*, 1996). It is suggested that this effect is especially salient when the target speech stream is masked by background noise.

While there are evident effects of hearing loss throughout central auditory pathways, as demonstrated in animal models, SNHL may selectively affect the encoding of different elements of speech. Two acoustic aspects of the signal, the temporal envelope and temporal fine structure (TFS), are pertinent for addressing the communication needs of individuals with SNHL. The envelope arises from the relatively slow fluctuations in the overall amplitude of the signal and contributes to the sensation of loudness (Rennies *et al.*, 2010). The TFS refers to the relatively faster fluctuations of sound pressure that “carry” the envelope and contain the signal’s spectral structure; TFS cues convey the timbre or quality of a signal (Moore, 2008).

The advent of cochlear implants (CIs), which at present deliver the envelope of the stimulus but not the TFS (Rubinstein, 2004), has provided a new understanding of the difficulties encountered by listeners with SNHL. In quiet, speech understanding can be achieved with as little as four channels of envelope coding, but CI users continue to have impaired music and speech-in-noise perception, presumably

<sup>a)</sup>Present address: Department of Hearing and Speech Sciences, 0100 LeFrak Hall, College Park, MD 20742.

<sup>b)</sup>Author to whom correspondence should be addressed. Also at: Department of Neurobiology and Physiology, 2205 Tech Drive Hogan 2-160, Evanston, IL 60208. Electronic mail: [nkraus@northwestern.edu](mailto:nkraus@northwestern.edu)

because CIs do not provide TFS cues (Strelcyk and Dau, 2009; Hopkins and Moore, 2011; Papakonstantinou *et al.*, 2011; Limb and Rubinstein, 2012). TFS cues are not sufficient, however; in fact, they contribute to perception of speech in noise only in the presence of the envelope (Swaminathan and Heinz, 2012). To achieve acceptable speech intelligibility in a variety of contexts, then, listeners must be able to process both aspects of speech signals.

Individuals with mild-to-moderate degrees of SNHL (i.e., who would not be candidates for implantation) also have difficulty hearing in noise (Divenyi and Haupt, 1997). Perceptual experiments have confirmed that individuals with SNHL do not have deficits in envelope encoding; in fact, in some cases they may have enhanced envelope encoding as demonstrated by better than normal amplitude modulation detection thresholds (Füllgrabe *et al.*, 2003) or gap detection thresholds (Horwitz *et al.*, 2011). Mixed results, however, have been obtained in experiments that have evaluated processing of TFS in individuals with normal hearing and SNHL. There is evidence of TFS encoding deficits in younger adults with SNHL who have been matched for age with normal hearing counterparts (Ardoint *et al.*, 2010). Age-related reductions in TFS have been found in perceptual (Grose and Mamo, 2010) and neurophysiological studies (Anderson *et al.*, 2012). It is therefore important to study the processing of speech cues in individuals with SNHL compared to age-matched, normal hearing controls.

The relative strength of encoding of envelope and TFS cues has been tested neurophysiologically. Chinchillas with noise-induced SNHL have abnormally enhanced coding of the stimulus envelope in auditory nerve fibers but the ability to encode TFS is unaffected when stimuli are presented in quiet (Kale and Heinz, 2010). A follow-up study revealed that the hearing-impaired chinchilla's ability to encode TFS is degraded substantially more in noise backgrounds than in quiet (Henry and Heinz, 2012). Because of this deficit, the salience of TFS cues may be reduced by excessive encoding of the ENV, reducing the listener's ability to make use of the TFS cues that are important for hearing in noise.

To increase understanding of the communication difficulties encountered by the hearing impaired listener, here the effects of hearing loss were investigated in the human brainstem using far-field electrophysiological techniques. The brainstem's frequency following response (FFR) was chosen because by eliciting the response to speech sounds presented in alternating polarities, the representation of the envelope and TFS can be assessed independently (Aiken and Picton, 2008; Campbell *et al.*, 2012). Furthermore, animal models of SNHL have found alterations in the balance of excitatory and inhibitory neurotransmitters in the inferior colliculus (Vale and Sanes, 2002), a putative primary generator of the FFR (Chandrasekaran and Kraus, 2010). The FFR therefore provides an ideal model to understand the converging peripheral and central effects of SNHL.

Aging affects the representation of both the speech envelope and TFS in normal hearing listeners (Anderson *et al.*, 2012); therefore, the normal-hearing (NH) and hearing-impaired (HI) participants were matched on age. Based on animal models of neurophysiologic encoding and human

perceptual experiments, it is hypothesized that individuals with SNHL have excessive encoding of the envelope in the FFR of the brainstem compared to that of individuals with normal hearing. To test this hypothesis group differences were assessed in the subcortical representation of a 40 ms [da] syllable, elicited in two conditions. In the first stimulus condition, responses were recorded to an unamplified stimulus in both the NH and the HI groups; in the second stimulus condition the same unamplified NH responses were compared to those of a customized, amplified stimulus in the HI group, in an effort to equate for audibility in cases of hearing loss.

## II. METHODS

### A. Participants

Thirty adult participants were recruited from the Chicago area; 15 with normal hearing [ages 61 to 68 years, mean 64.07, standard deviation (SD) 2.09] and 15 with hearing loss (60 to 71, mean 64.07, SD 3.39). Participants in the normal hearing (NH) group had clinically normal hearing, defined as: air conduction thresholds  $\leq 20$  dB hearing level (HL) from 125–4000 Hz bilaterally,  $\leq 40$  dB HL at 6000 and 8000 Hz, and pure-tone averages (PTA; average from 500–4000 Hz)  $\leq 15$  dB HL. Participants in the hearing impaired (HI) group had hearing levels ranging from mild-to-moderate hearing loss, defined as thresholds from 125 to 2000 Hz  $< 40$  dB HL and from 3000 to 8000 Hz  $< 70$  dB HL and PTA  $\geq 30$  dB HL and  $< 45$  dB HL. All subjects had normal click-evoked auditory brainstem response (ABR) latencies (wave V  $< 6.8$  ms), measured by a 100- $\mu$ s click stimulus presented at 80 dB SPL (peak equivalent) at a rate of 31.4 Hz. In addition, there was no evidence of a conductive hearing loss (no air-bone gaps  $\geq 10$  dB HL at two or more frequencies) or inter-aural asymmetry (defined as  $\geq 15$  dB HL audiometric difference at two or more frequencies or an inter-aural click-ABR difference  $> 0.2$  ms). No participant had ever worn hearing aids or reported a history of neurologic conditions. See Fig. 1 for average group hearing thresholds and representative click-ABRs.

The groups were sex-matched [ $\chi^2$  (chi square):  $p > 0.1$ ]—the NH and HI groups included five and seven males, respectively. Participants had normal intelligence quotient (IQ) scores ( $> 85$  on the Wechsler Abbreviated Scale of Intelligence, Psychological Corp, San Antonio, TX) and were matched for IQ. See Table I for means and SDs of group characteristics. All procedures were approved by the Institutional Review Board of Northwestern University. Participants provided informed consent and were compensated for their time.

### B. Electrophysiology

#### 1. Stimulus

The stimulus was a 40-ms [da] syllable synthesized in a Klatt-based synthesizer (Klatt, 1980). The stimulus began with an onset burst that was followed by the consonant-to-vowel (CV) transition. Although the stimulus does not include the steady-state vowel, it is perceived as a CV syllable. The

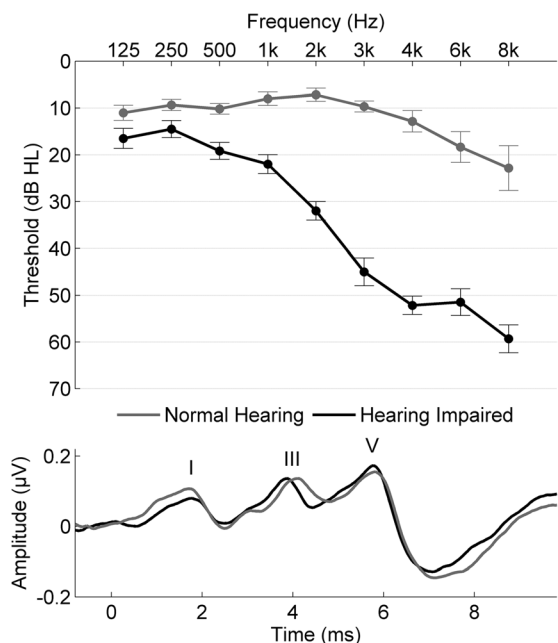


FIG. 1. Top: Average pure-tone thresholds for the normal hearing group (gray) and the hearing impaired group (black). Bottom: Response waveforms to the click stimulus from individuals in the normal hearing group and the hearing impaired group. An apparent delay at wave I in the hearing impaired participant is no longer present at wave V.

[da] syllable was chosen because the perception of stop consonants can be challenging for both young and old listeners (Miller and Nicely, 1955), providing a more ecologically valid stimulus than the clicks or tonebursts that are typically used in ABR assessments. The [da] stimulus elicits a robust FFR across the lifespan, and its brainstem encoding reflects supra-threshold central auditory processing and its behavioral correlates (Anderson *et al.*, 2011). After an initial 5-ms onset burst, the fundamental frequency ( $F_0$ ) of the stimulus rose linearly from 103 to 125 Hz while the formants shifted as follows:  $F_1$ : 220  $\rightarrow$  720 Hz;  $F_2$ : 1700  $\rightarrow$  1240 Hz;  $F_3$ : 2580  $\rightarrow$  2500 Hz. The fourth (3600 Hz) and fifth (4500 Hz) formants remained constant for the duration of the stimulus. A spectrogram and Fourier transform of the stimulus waveform are presented in Figs. 2(A) and 2(B), respectively.

The [da] stimulus was presented binaurally via the Biologic Navigator Pro System (Natus Medical, Inc., Mundelein, IL) at a rate of 10.9 Hz (inter-stimulus interval of 52 ms) through electromagnetically shielded insert earphones (ER-3A, Etymotic Research, Elk Grove Village, IL) at 80.3 dB sound pressure level (SPL) in quiet and noise

TABLE I. Participants are matched on age for listed characteristics (means, with SDs, reported; all  $p$ 's  $>$  0.1, except for wave I). Click I and click V refer to latency values for wave components.

	Normal Hearing ( $N = 15$ )	Hearing Impaired ( $N = 15$ )
Sex	Males = 5	Males = 7
Age	64.07 (2.09)	64.07 (3.39)
IQ	115.60 (15.45)	117.73 (14.04)
Click I	1.71 ms (.29)	1.94 ms (.19) <sup>a</sup>
Click V	6.03 ms (.32)	6.07 ms (.34)

<sup>a</sup> $p = 0.03$ .

conditions. Prior to each recording, the click and [da] stimuli were calibrated to 80 dB SPL using a Brüel and Kjær 2238 Mediator sound level meter coupled to an insert earphone adaptor. The SPL for each stimulus was sampled over 60 s to obtain the average SPL. In quiet, only the [da] was presented; in noise, a background of pink noise was also played at +10 dB signal-to-noise ratio (SNR). During the recording session, the participants watched muted movies of their choice with subtitles to facilitate an awake but quiet state.

The [da] stimulus was presented in alternating polarities, allowing for the creation of responses consisting of both the sum and the difference of the two presentation polarities (Campbell *et al.*, 2012). With the summed responses, the non-inverting envelope component of the response is enhanced while the inverting TFS component is minimized. On the other hand, with the subtracted responses, the inverting TFS component is enhanced while the envelope component is substantially reduced (Aiken and Picton, 2008). Because 3000 repetitions were averaged, it was expected that the effects of random pink noise would be equivalent for added and subtracted polarities. Figure 2(C) and 2(D) display the averaged NH response to added (envelope) and subtracted (TFS) polarities.

## 2. Amplification in cases of SNHL

To equate for effects of peripheral hearing loss, the [da] stimulus was modified for individuals in the HI group using the National Acoustics Laboratory-Revised algorithm (NAL-R) (Byrne and Dillon, 1986), with a custom program in MATLAB (The MathWorks, Inc., Natick, MA) based on individual hearing thresholds. The amplified presentation intensities did not exceed 90 dB SPL or any individual's loudness discomfort levels. For participants with hearing loss, both the amplified and the unamplified [da] stimuli were presented in quiet and noise for a total of four stimulus conditions (as opposed to the NH group, who were only presented with two stimulus conditions). The order of presentation conditions was held constant to minimize order effects (i.e., fatigue) on group differences. In the NH group the quiet condition was followed by the noise condition. In the HI group, the order was as follows: (1) unamplified stimulus presented in quiet, (2) amplified stimulus presented in quiet, (3) unamplified stimulus presented in noise, and (4) amplified stimulus presented in noise.

## 3. Recording

A vertical montage of four Ag-AgCl electrodes (Cz active, Fpz ground, earlobe references) was used with all contact impedances  $<$  5 k $\Omega$ . Responses were recorded using the Biologic Navigator Pro System. A criterion of  $\pm 23 \mu\text{V}$  was used for online artifact rejection. Two blocks of 3000 artifact-free sweeps were collected in each condition for each participant and averaged using an 85.3-ms window, including a 15.8-ms prestimulus period. The responses were sampled at 12 kHz and were online bandpass filtered from 100 to 2000 Hz (Butterworth filter, 12 dB/octave, zero phase-shift) to minimize disruption by the low-frequency cortical response and to sample energy up to the phase-locking limits of the brainstem (Chandrasekaran and Kraus, 2010).



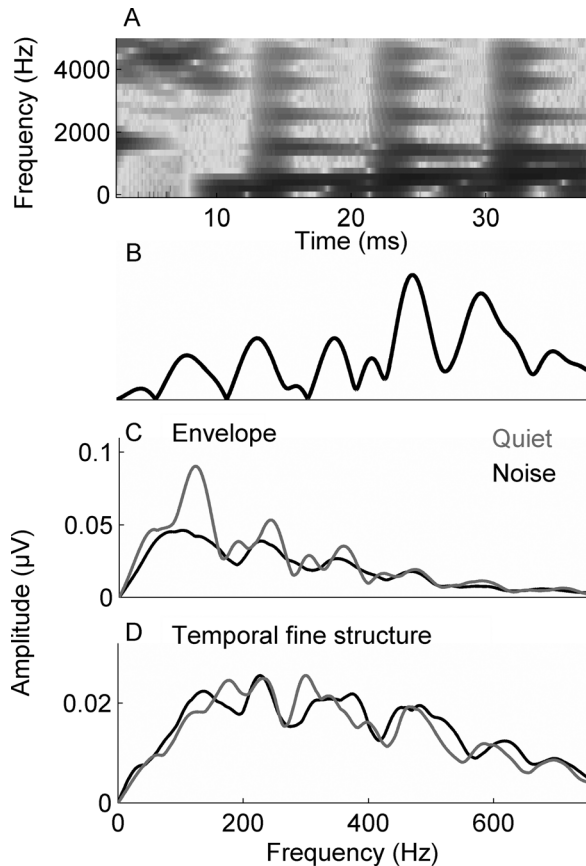


FIG. 2. (A) The spectrogram of the 40-ms syllable [da]. Fast Fourier transforms were calculated from 20–42 ms for the stimulus (B) and in responses to the envelope (C) and the TFS (D) in quiet and in noise. The average responses for the NH group are displayed.

#### 4. Data analysis

*a. Spectral amplitudes.* To obtain an index of frequency-specific phase-locked activity in the evoked response, spectral amplitudes were computed using fast Fourier transforms (FFTs) on the envelope (added, indicated henceforth as ADD) and TFS (subtracted, indicated henceforth as SUB) responses. Zero padding was applied prior to the transform to increase the resolution of the spectral display to 1 Hz/point, and the FFTs were run with a Hanning window and a 4 ms ramp. FFTs were run on a time window of 20–42 ms (60-Hz bins), the time range that encompasses the FFR in older adults based on visual inspection of the periodic components of the brainstem response waveform. The temporal envelope was calculated by determining the amplitude of the envelope-dominated low frequencies—the fundamental frequency ( $F_{0-ADD}$ ) and second ( $H_{2-ADD}$ ) and third ( $H_{3-ADD}$ ) harmonics (three values), where ADD indicates FFT amplitudes from the FFR to added polarities; the TFS was calculated by determining the amplitude of the TFS-dominated harmonics of the first formant, fourth ( $H_{4-SUB}$ ), fifth ( $H_{5-SUB}$ ), and sixth ( $H_{6-SUB}$ ) harmonic (three values), where SUB indicates FFT amplitudes from the FFR to subtracted polarities.

A measure of relative envelope and TFS representation was computed for a given harmonic,  $X$ , as follows:

$$\Delta H_X = H_{X-ADD} - H_{X-SUB},$$

where  $H_X$  refers to the FFT amplitude of each harmonic from  $F_0$  to  $H_6$ , ADD refers to spectral amplitudes of responses to added inverting polarities (emphasizing envelope), and SUB refers to amplitudes when polarities are subtracted (emphasizing TFS). For example, the difference in amplitude between  $H_3$  in the envelope- vs TFS-following responses would be computed as  $\Delta H_3 = H_{3-ADD} - H_{3-SUB}$ .

*b. Statistical analyses.* All statistical analyses were conducted in SPSS, version 18.0 (SPSS, Inc., Chicago, IL). Multivariate analysis of variance (MANOVA) was used to compare the amplitudes of envelope ( $F_{0-ADD}$ – $H_{3-ADD}$ ) and TFS components ( $H_{4-SUB}$ – $H_{6-SUB}$ ) of the [da] syllable between the NH and HI groups. A MANOVA was also used to compare the  $\Delta H_X$  across the frequency range of  $F_0$ – $H_6$ . These analyses were completed for two stimulus conditions. In condition 1, both groups received unamplified stimuli at 80.3 dB SPL. In condition 2, the HI group received an amplified stimulus tailored to their hearing profile (see below). Normal distribution and homogeneity of variance was verified by the Shapiro-Wilk and Levene’s test, respectively. Bonferroni corrections for multiple comparisons were applied as needed, and all  $p$  values reflect two-tailed tests.

### III. RESULTS

#### A. Summary of results

In stimulus condition 1, the unamplified [da] stimulus was presented to both the NH and HI groups. In stimulus condition 2, the unamplified [da] stimulus was presented to the NH group but individually amplified [da] stimuli were presented to participants in the HI group. Overall results indicated that individuals with hearing loss have an imbalance in the representation of the envelope and TFS components of speech compared to individuals with normal hearing [see Table II for the amplitudes of the  $F_0$  and its harmonics in the brainstem FFR to the envelope and TFS of the [da] and Table III for differences in amplitude ( $\Delta H_X$ )]. This finding is present when the speech syllable is presented in quiet at a moderately loud level (80 dB SPL) or at a fixed +10 SNR. In individuals with hearing loss, the representation of the envelope was disproportionately high relative to TFS representation based on envelope-to-TFS differences in amplitude. This resulted in an effective “relative deficit” in TFS representation.

#### B. Stimulus condition 1—Unamplified [da] stimulus

- (1) Envelope. In the HI group, the amplitude of the envelope ( $F_{0-ADD}$ – $H_{3-ADD}$ ) was larger than in the NH group in response to the [da] stimulus presented in a background of pink noise [ $F(1,26) = 3.796$ ,  $p = 0.022$ ], but not to the [da] presented in quiet [ $F(1,26) = 1.273$ ,  $p = 0.304$ ] [Figs. 3(A) and 3(B)].
- (2) TFS. No differences were noted in the amplitude of the TFS ( $H_{4-SUB}$ – $H_{6-SUB}$ ) between the HI and NH groups in either quiet [ $F(1,26) = 2.391$ ,  $p = 0.092$ ] or noise [ $F(1,26) = 0.742$ ,  $p = 0.549$ ] [Figs. 3(C) and 3(D)].

TABLE II. Mean spectral amplitudes, with standard deviations, of the brainstem FFR to the [da] stimulus presented in quiet and noise. Amplification is applied in cases of hearing loss.

	Condition 1. Unamplified [da] stimulus		Condition 2. Amplified [da] stimulus
	Normal hearing mean amplitude ( $\mu\text{V}$ ) (S.D.)	Hearing impaired mean amplitude ( $\mu\text{V}$ ) (S.D.)	Hearing impaired mean amplitude ( $\mu\text{V}$ ) (S.D.)
Envelope (spectral FFT amplitudes of the FFR to added polarities)			
Quiet			
$F_{0\text{-ADD}}$	0.074 (0.029)	0.079 (0.044)	0.083 (0.056)
$H_{2\text{-ADD}}$	0.043 (0.034)	0.056 (0.041)	0.074 (0.039)
$H_{3\text{-ADD}}$	0.030 (0.017)	0.043 (0.026)	0.045 (0.020)
Noise			
$F_{0\text{-ADD}}$	0.075 (0.029)	0.078 (0.057)	0.080 (0.045)
$H_{2\text{-ADD}}$	0.043 (0.033)	0.072 (0.040)	0.082 (0.037)
$H_{3\text{-ADD}}$	0.030 (0.017)	0.047 (0.021)	0.047 (0.018)
TFS (spectral FFT amplitudes of the FFR to subtracted polarities)			
Quiet			
$H_{4\text{-SUB}}$	0.019 (0.015)	0.022 (0.012)	0.018 (0.013)
$H_{5\text{-SUB}}$	0.011 (0.018)	0.012 (0.007)	0.012 (0.010)
$H_{6\text{-SUB}}$	0.008 (0.005)	0.006 (0.005)	0.008 (0.007)
Noise			
$H_{4\text{-SUB}}$	0.020 (0.019)	0.022 (0.014)	0.018 (0.013)
$H_{5\text{-SUB}}$	0.010 (0.007)	0.014 (0.010)	0.012 (0.010)
$H_{6\text{-SUB}}$	0.007 (0.005)	0.009 (0.008)	0.008 (0.007)

(3) Spectral differences between FFR FFTs to added and subtracted polarities of the stimulus. The amplitude differences between envelope and TFS representation were significantly larger in the HI group compared to the NH group in noise [ $F(1,23) = 3.506$ ,  $p = 0.013$ ] but not in quiet [ $F(1,23) = 1.712$ ,  $p = 0.163$ ] [Fig. 3(I)].

### C. Stimulus condition 2—Amplification applied in the HI group

(1) Envelope. In the HI group, the amplitude of the envelope-dominated frequencies ( $F_{0\text{-ADD}}-H_{3\text{-ADD}}$ ) was

larger than in the NH group in the FFR to the [da] stimulus presented in both noise and quiet backgrounds [noise:  $F(3,26) = 4.623$ ,  $p = 0.010$ ; quiet:  $F(3,26) = 3.042$ ,  $p = 0.047$ ] [Figs. 3(E) and 3(F)].

(2) TFS. No differences were found in the amplitude of the TFS ( $H_{4\text{-SUB}}-H_{6\text{-SUB}}$ ) between the HI and NH groups in noise [ $F(3,26) = 0.552$ ,  $p = 0.671$ ] or in quiet [ $F(3,26) = 0.151$ ,  $p = 0.928$ ] [Figs. 3(G) and 3(H)].

(3) Spectral differences between FFR FFTs to added and subtracted polarities of the stimulus. The amplitude differences between envelope and TFS representation were significantly larger in the HI group than in the NH group

TABLE III. Mean differences in spectral amplitudes (envelope – TFS) in NH and HI groups, indicating significant differences between the hearing-impaired groups (unamplified and amplified) and the normal-hearing group.

	Unamplified [da] stimulus		Amplified [da] stimulus
	Normal Hearing Mean $\Delta H_x$ in $\mu\text{V}$ (S.D.)	Hearing Impaired Mean $\Delta H_x$ in $\mu\text{V}$ (S.D.)	Hearing Impaired Mean $\Delta H_x$ in $\mu\text{V}$ (S.D.)
Quiet			
$\Delta F_0$	0.059 (0.027)	0.063 (0.048)	0.063 (0.058)
$\Delta H_2$	0.017 (0.030)	0.039 (0.038)	0.055 (0.038) <sup>b</sup>
$\Delta H_3$	0.009 (0.017)	0.028 (0.027) <sup>a</sup>	0.030 (0.016) <sup>b</sup>
$\Delta H_4$	-0.001 (0.015)	0.010 (0.022)	0.025 (0.014) <sup>c</sup>
$\Delta H_5$	0.000 (0.008)	0.009 (0.011) <sup>a</sup>	0.013 (0.009) <sup>c</sup>
$\Delta H_6$	-0.002 (0.006)	0.036 (0.007) <sup>a</sup>	0.002 (0.009)
Noise			
$\Delta F_0$	0.055 (0.032)	0.057 (0.059)	0.060 (0.051)
$\Delta H_2$	0.019 (0.035)	0.054 (0.039) <sup>a</sup>	0.063 (0.040) <sup>b</sup>
$\Delta H_3$	0.008 (0.017)	0.024 (0.016) <sup>a</sup>	0.032 (0.017) <sup>b</sup>
$\Delta H_4$	-0.002 (0.019)	0.018 (0.015) <sup>b</sup>	0.020 (0.011) <sup>b</sup>
$\Delta H_5$	0.001 (0.010)	0.007 (0.011)	0.012 (0.009) <sup>b</sup>
$\Delta H_6$	0.000 (0.010)	0.001 (0.005)	0.004 (0.009)

<sup>a</sup> $p < 0.05$ .

<sup>b</sup> $p < 0.01$

<sup>c</sup> $p < 0.001$ .

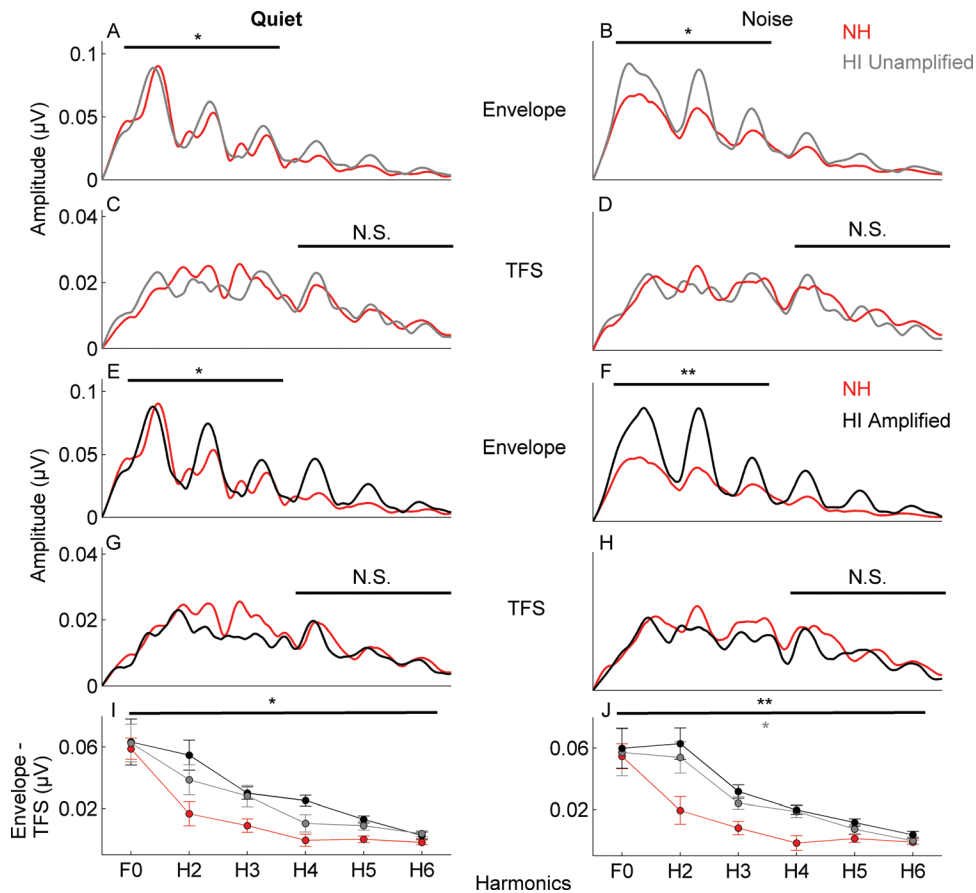


FIG. 3. Comparison of response spectra to the 40-ms stimulus over the frequency following response (20–42 ms) for the NH (red) versus HI groups (unamplified stimulus: gray; amplified stimulus: black) in quiet and in noise. (A), (B) Response spectra to the envelope. The HI group (in response to the unamplified stimulus) has higher amplitudes in the envelope-dominated low frequencies ( $F_0$ – $H_3$ ) in noise. (C), (D) There are no group differences in the response to the TFS-dominated high frequencies ( $H_4$ – $H_6$ ). (E), (F) In response to the envelope, the HI group (amplified stimulus) has higher amplitudes in the low frequencies (overall) in quiet and noise. (G), (H) There are no group differences in response to the TFS between the HI (amplified stimulus) and NH groups. (I), (J) Larger delta  $H_x$  differences over the range of  $F_0$ – $H_6$  are noted in the quiet condition for the HI group (amplified stimulus) compared to the NH group and in the noise condition for the HI group in response to both unamplified and amplified stimuli. Error bars in (I) and (J): 1 S.E. \* $p < 0.05$ , \*\* $p < 0.01$ .

in quiet and in noise [noise:  $F(1,23) = 5.403$ ,  $p = 0.001$ ; quiet:  $F(1,23) = 4.221$ ,  $p = 0.005$ ] [Fig. 3(J)].

#### IV. DISCUSSION

Far-field electrophysiology was used to measure subcortical representation of speech cues in older adults with and without SNHL. These results demonstrate greater spectral representation in response to the temporal envelope in the SNHL group, with equivalent representation of the TFS between groups in response to an 80 dB [da] stimulus presented in noise at a fixed +10 SNR. The amplitude differences between envelope and TFS representation were greater in the HI group, consistent with the suggestion of Kale and Heinz (2010) that the relative coding of TFS and envelope may be disrupted in the case of SNHL, thereby providing a possible neurophysiologic basis for the difficulties encountered by listeners with SNHL when trying to understand speech in background noise.

There are a number of potential mechanisms responsible for the boosting of envelope-following brainstem responses in individuals with hearing loss. The levels of neurotransmitters in the inferior colliculus (IC) may be disrupted, such that the usual balance of excitation and inhibition is altered with a net increase in volume. While this hypothesis cannot be directly tested in humans, there is support from animal models. For example, deafness results in larger and longer excitatory post synaptic current amplitudes and durations, along with reduced inhibitory strength in IC neurons (Vale and Sanes, 2002). The corticofugal pathway extends all the

way to the outer hair cells of the inner ear, through activation of efferent fibers in the medial olivocochlear bundle, modulating the gain provided by the outer hair cells (Perrot *et al.*, 2006). The corticofugal pathway can affect the response properties of neurons in the IC based on experience (Gao and Suga, 2000); therefore, neurons in the IC may be more excitable due to a deprivation of auditory input and the resulting decrease in gain modulation by inhibitory neurotransmission. This decreased inhibition, along with increased excitability throughout the central auditory pathway, may result in a distortion of speech cue representation in the IC as indexed by the FFR.

An alternate explanation for the findings may be that the enhanced subcortical envelope encoding is the result of greater broadband activity. Individuals with hearing loss have broader tuning curve bandwidths (Florentine *et al.*, 1980); therefore, greater broadband activation may result in higher spectral amplitudes. In the Kale and Heinz study, however, no one-to-one correspondence between greater bandwidths and greater envelope coding was found, suggesting that other mechanisms contribute to this phenomenon (Kale and Heinz, 2010).

Behavioral studies have shown that TFS can be reduced with hearing loss when frequency selectivity is not impaired (Lorenzi *et al.*, 2009; Strelcyk and Dau, 2009). Reduced TFS in the presence of normal frequency selectivity in the auditory nerve has also been found in an animal model (Heinz and Swaminathan, 2009). In the current study, equivalent TFS representation in normal hearing and hearing impaired participants may be due to using one presentation level and

one SNR in the background noise stimulus condition. Therefore, temporal coding may not have been sufficiently disrupted to obtain a noise-induced effect similar to that of Henry and Heinz (2012). Furthermore, the lack of differences in TFS encoding may be a reflection of a limited or reduced range of TFS values in older adults rather than a lack of difference between individuals with normal hearing and with hearing impairment. Further parametric investigations comparing TFS encoding in age-matched young adults with and without hearing loss is warranted to address this concern. Future work should compare responses at different SNRs and presentation levels and should include a behavioral measure of frequency selectivity, such as psychophysical tuning curves, to better understand the underlying mechanisms driving the imbalance in envelope and TFS representation.

### A. Amplified versus unamplified stimuli

In a previous study, the NAL-R (Byrne and Dillon, 1986) algorithm was used to partially equate for hearing loss (Anderson *et al.*, 2011). The same amplification procedure was used in this study and similar group differences were found for both conditions, with the HI group having greater envelope encoding and larger envelope-to-TFS ratios. Because of the pronounced differences between the NH and HI groups when testing the HI group with amplified stimuli, it was necessary to ensure that the differences were not driven by presenting higher-level (i.e., greater intensity) stimuli to the HI group. Therefore, an unamplified stimulus was used to compare responses between the NH and HI groups. Even with this unamplified stimulus, enhanced encoding of the envelope was maintained in the HI group. Taken together, the findings of enhanced encoding of the envelope to an unamplified stimulus and essentially equal response amplitudes to the TFS of both amplified and unamplified stimuli indicate that amplification was not a confounding factor in the current results.

### B. Relative encoding of envelope versus temporal fine structure

Consistent with the work of Kale and Heinz (2010), no differences were found for the absolute representation of TFS; rather, the differences between envelope and TFS representation were greater in the HI group, resulting in a *relative* deficit of TFS. Importantly, this deficit is demonstrated in human older adults with age-related SNHL; therefore, the findings in an animal model of noise-induced SNHL can indeed be extended to humans who have experienced a gradual loss of peripheral hearing. Given the demonstrated importance of TFS for understanding speech in noise (Gnansia *et al.*, 2009; Hopkins and Moore, 2011; Papakonstantinou *et al.*, 2011), the relative deficit in TFS coding may play a critical role in the difficulties encountered by older adults with SNHL when trying to listen in noise.

The role of TFS in speech perception has been a controversial subject partly due to differences in experimental protocols. Normal hearing individuals have improved perception when listening to a signal in fluctuating background noise

compared to listening to a signal in steady noise, an improvement known as masking release (Füllgrabe *et al.*, 2006). Hearing impaired individuals, however, do not experience this masking release in fluctuating noise conditions (Duquesnoy, 1983). It has been hypothesized that a deficit in TFS representation prevents hearing impaired individuals from processing or hearing speech information in the dips of a fluctuating masker, because the auditory system uses TFS information to determine if the target signal is present in the dips (Lorenzi *et al.*, 2006). Bernstein and Brungart (2011), however, argue that the lack of masking release in SNHL in these studies may be attributed to experimental protocol rather than to effects of hearing loss. For example, to control for individual differences in audibility, hearing impaired individuals are typically tested at higher SNRs than those used for normal hearing individuals, and masking release decreases with higher SNRs (Bernstein and Grant, 2009). Therefore, the reduced masking release found in the hearing impaired may be attributed to the higher SNRs rather than to a deficit in TFS representation.

As mentioned previously, age may have been a confound in some reports of decreased encoding of TFS in individuals with SNHL (Lorenzi *et al.*, 2009), but recent work has demonstrated reduced TFS encoding in hearing impaired young adult listeners compared to an age-matched group of normal hearing listeners (Ardoint *et al.*, 2010). Nevertheless, when investigating the encoding of TFS or envelope, it is important to consider the age group being studied. Ruggles *et al.* (2012) found that middle-aged adults rely on TFS cues to selectively attend to a target stream from competing masker streams, whereas young adults rely on envelope cues for this same task. Therefore, group differences in levels of envelope versus TFS may change depending on the age group being studied. Here, these results show that the *relative* encoding of TFS is reduced in older adults (ages 60–71) with hearing loss compared to age-matched older adults with normal hearing. Future work will directly explore the role of relative envelope-TFS representation for speech perception in noise.

### C. Clinical implications

These results have ramifications for the management of individuals with hearing loss, both for the development of hearing aid and CI processing algorithms and for the creation of auditory training programs. A hearing aid algorithm that provides audibility and enhanced envelope encoding at the expense of TFS will result in improved hearing in quiet situations; however, in noise the speech understanding provided by amplification may be no better or even worse than no amplification because of an alteration in the balance between envelope and TFS coding. Bilateral hearing aid users themselves attest to this; in the first author's clinical experience, they often remove one hearing aid in noise to reduce overall loudness and to "hear better," even when their hearing aids provide advanced multidirectional processing to enhance the signal-to-noise ratio. The demonstration of envelope and TFS differences in the FFR of the brainstem suggests a critical role for electrophysiology in the assessment of hearing aid algorithms in both research and clinical



settings. For example, brainstem responses can be used to objectively compare envelope and TFS representation under multiple amplification procedures.

## D. Conclusions

These results demonstrate that older adults with hearing loss have a relative deficit of TFS encoding in the FFR, as demonstrated by greater encoding of the envelope without a similar enhancement of TFS. These results are similar to those of [Kale and Heinz \(2010\)](#), who demonstrated this deficit in the auditory nerves of chinchillas with induced SNHL. The presumed result of this misbalance is that the envelope cues overpower, or “swamp out,” the response, such that the details are not discernible (because the TFS cues become less salient). This effect is exacerbated in noise—the very environment in which TFS cues are critical for accurate speech perception. Because the current study’s methodology employs far-field recordings to evaluate neural responses in humans as opposed to the direct recordings of neural firing or cellular function performed in animals, it is not possible to directly test mechanisms responsible for this deficit, such as changes in the levels of neurotransmitters. Future work should consider interactions of peripheral and central input. For example, an input-output function model might determine whether or not this deficit reflects a loss of the compression that would normally be provided by functioning outer hair cells.

Management of individuals with hearing loss should include consideration of ways to boost access to TFS cues. There is interest in providing access to TFS cues to CI users, and one CI manufacturer has attempted to deliver TFS cues in the low frequencies up to 1000 Hz through changes in stimulation rate ([Arnoldner et al., 2007](#); [Riss et al., 2009](#); [Galindo et al., 2012](#)). In addition, there is evidence in CI users of benefits for both long-term ([Gfeller et al., 2000](#)) and short-term ([Galvin et al., 2009](#)) musical training on tasks that require perception of TFS, such as melodic contour perception, demonstrating that plasticity of the auditory system is possible, even in the presence of impoverished auditory cues. Musical training also confers advantages for processing speech in noise in older adults with normal hearing ([Parbery-Clark et al., 2011](#); [Parbery-Clark et al., 2012](#)). Short-term computer-based auditory training improves speech understanding in noise and neural timing in older adults with and without hearing loss ([Anderson et al., 2013](#)). Listeners with hearing loss, then, may benefit from auditory training protocols which train them to re-weight TFS cues relative to the envelope. The FFR may one day provide an index of this re-weighting.

## ACKNOWLEDGMENTS

The authors thank the participants who participated in this study, Trent Nicol, Jen Krizman, and Adam Tierney for their comments on the manuscript, and Erica Knowles for her help with data analysis. This work was supported by the National Institutes of Health (Grant Nos. T32 DC009399-01A10 and RO1 DC10016) and the Knowles Hearing Center.

- Aiken, S. J., and Picton, T. W. (2008). “Envelope and spectral frequency-following responses to vowel sounds,” *Hear. Res.* **245**, 35–47.
- Aizawa, N., and Eggermont, J. (2006). “Effects of noise-induced hearing loss at young age on voice onset time and gap-in-noise representations in adult cat primary auditory cortex,” *J. Assoc. Res. Otolaryngol.* **7**, 71–81.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). “Aging affects neural precision of speech encoding,” *J. Neurosci.* **32**, 14156–14164.
- Anderson, S., Parbery-Clark, A., Yi, H.-G., and Kraus, N. (2011). “A neural basis of speech-in-noise perception in older adults,” *Ear Hear.* **32**, 750–757.
- Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). “Reversal of age-related neural timing delays with training,” *Proc. Natl. Acad. Sci. USA* **110**, 4357–4362.
- Ardoint, M., Sheft, S., Fleuriet, P., Garnier, S., and Lorenzi, C. (2010). “Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss,” *Int. J. Audiol.* **49**, 823–831.
- Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W. D., and Hamzavi, J. S. (2007). “Speech and music perception with the new fine structure speech coding strategy: Preliminary results,” *Acta Oto-Laryngol.* **127**, 1298–1303.
- Bernstein, J. G. W., and Brungart, D. S. (2011). “Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio,” *J. Acoust. Soc. Am.* **130**, 473–488.
- Bernstein, J. G., and Grant, K. W. (2009). “Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners,” *J. Acoust. Soc. Am.* **125**, 3358.
- Byrne, D., and Dillon, H. (1986). “The National Acoustic Laboratories’ (NAL) new procedure for selecting the gain and frequency response of a hearing aid,” *Ear Hear.* **7**, 257–265.
- Campbell, T., Kerlin, J., Bishop, C., and Miller, L. (2012). “Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography,” *Ear Hear.* **33**, 144–150.
- Chandrasekaran, B., and Kraus, N. (2010). “The scalp-recorded brainstem response to speech: Neural origins and plasticity,” *Psychophysiol.* **47**, 236–246.
- Dallos, P., Zheng, J., and Cheatham, M. A. (2006). “Prestin and the cochlear amplifier,” *J. Physiol.* **576**, 37–42.
- Divenyi, P. L., and Haupt, K. M. (1997). “Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation,” *Ear Hear.* **18**, 189–201.
- Dong, S., Rodger, J., Mulders, W. H. A. M., and Robertson, D. (2010). “Tonotopic changes in GABA receptor expression in guinea pig inferior colliculus after partial unilateral hearing loss,” *Brain Res.* **1342**, 24–32.
- Duquesnoy, A. (1983). “Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons,” *J. Acoust. Soc. Am.* **74**, 739–743.
- Florentine, M., Buus, S., Scharf, B., and Zwicker, E. (1980). “Frequency selectivity in normally-hearing and hearing-impaired observers,” *J. Speech Hear. Res.* **23**, 646–669.
- Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). “Masking release for consonant features in temporally fluctuating background noise,” *Hearing Research* **211**, 74–84.
- Füllgrabe, C., Meyer, B., and Lorenzi, C. (2003). “Effect of cochlear damage on the detection of complex temporal envelopes,” *Hear. Res.* **178**, 35–43.
- Galindo, J., Lassaletta, L., Mora, R. P., Castro, A., Bastarrica, M., and Gavilán, J. (2012). “Fine structure processing improves telephone speech perception in cochlear implant users,” *Eur. Arch. Otorhinolaryngol.* **270**, 1–7.
- Galvin, J. J., Fu, Q.-J., and Shannon, R. V. (2009). “Melodic contour identification and music perception by cochlear implant users,” *Ann. N.Y. Acad. Sci.* **1169**, 518–533.
- Gao, E., and 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.
- Gfeller, K. A., Christ, J. F., and Knutson, J. F. (2000). “Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients,” *J. Am. Acad. Audiol.* **11**, 390–406.
- Gnansia, D., Pean, V., Meyer, B., and Lorenzi, C. (2009). “Effects of spectral smearing and temporal fine structure degradation on speech masking release,” *J. Acoust. Soc. Am.* **125**, 4023–4033.



- Grose, J. H., and Mamo, S. K. (2010). "Processing of temporal fine structure as a function of age," *Ear Hear.* **31**, 755–760.
- Heinz, M., and Swaminathan, J. (2009). "Quantifying envelope and fine-structure coding in auditory nerve responses to chimaeric speech," *J. Assoc. Res. Otolaryngol.* **10**, 407–423.
- Henry, K. S., and Heinz, M. G. (2012). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," *Nat. Neurosci.* **15**, 1362–1364.
- Hopkins, K., and Moore, B. C. J. (2011). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," *J. Acoust. Soc. Am.* **130**, 334–349.
- Horwitz, A. R., Ahlstrom, J. B., and Dubno, J. R. (2011). "Level-dependent changes in detection of temporal gaps in noise markers by adults with normal and impaired hearing," *J. Acoust. Soc. Am.* **130**, 2928–2938.
- Jürgens, T., Kollmeier, B., Brand, T., and Ewert, S. D. (2011). "Assessment of auditory nonlinearity for listeners with different hearing losses using temporal masking and categorical loudness scaling," *Hear. Res.* **280**, 177–191.
- Kale, S., and Heinz, M. (2010). "Envelope coding in auditory nerve fibers following noise-induced hearing loss," *J. Assoc. Res. Otolaryngol.* **11**, 657–673.
- Klatt, D. (1980). "Software for a cascade/parallel formant synthesizer," *J. Acoust. Soc. Am.* **67**, 971–995.
- Kotak, V. C., Fujisawa, S., Lee, F. A., Karthikeyan, O., Aoki, C., and Sanes, D. H. (2005). "Hearing loss raises excitability in the auditory cortex," *J. Neurosci.* **25**, 3908–3918.
- Limb, C. J., and Rubinstein, J. T. (2012). "Current research on music perception in cochlear implant users," *Otolaryngol. Clinics N. Am.* **45**, 129–140.
- Lorenzi, C., Gilbert, G. T., Carn, H. L. S., Garnier, S. P., and Moore, B. C. J. (2006). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proc. Natl. Acad. Sci. USA* **103**, 18866–18869.
- Lorenzi, C., Debruelle, L., Garnier, S., Fleuriot, P., and Moore, B. C. J. (2009). "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal," *J. Acoust. Soc. Am.* **125**, 27–30.
- Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (2005). "Cochlear compression in listeners with moderate sensorineural hearing loss," *Hear. Res.* **205**, 172–183.
- Miller, G. A., and Nicely, P. E. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**, 338–352.
- Moore, B. C. J., Wojtczak, M., and Vickers, D. A. (1996). "Effect of loudness recruitment on the perception of amplitude modulation," *J. Acoust. Soc. Am.* **100**, 481–489.
- Moore, B. C. J. (2008). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," *J. Assoc. Res. Otolaryngol.* **9**, 399–406.
- Morita, T., Naito, Y., Nagamine, T., Fujiki, N., Shibasaki, H., and Ito, J. (2003). "Enhanced activation of the auditory cortex in patients with inner-ear hearing impairment: A magnetoencephalographic study," *Clin. Neurophysiol.* **114**, 851–859.
- Munro, K. J., and Blount, J. (2009). "Adaptive plasticity in brainstem of adult listeners following earplug-induced deprivation," *J. Acoust. Soc. Am.* **126**, 568–571.
- Papakonstantinou, A., Strelcyk, O., and Dau, T. (2011). "Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise," *Hear. Res.* **280**, 30–37.
- Perrot, X., Ryvlin, P., Isnard, J., Guénot, M., Catenoix, H., Fischer, C., Mauguière, F., and Collet, L. (2006). "Evidence for corticofugal modulation of peripheral auditory activity in humans," *Cereb. Cortex* **16**, 941–948.
- Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (2012). "Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults," *Front. Aging Neurosci.* **4**, 1–12.
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). "Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise," *Plos ONE* **6**, e18082.
- Rennies, J., Verhey, J. L., and Fastl, H. (2010). "Comparison of loudness models for time-varying sounds," *Acta Acust. Acust.* **96**, 383–396.
- Riss, D., Arnoldner, C., Reiß, S., Baumgartner, W. D., and Hamzavi, J. S. (2009). "1-year results using the opus speech processor with the fine structure speech coding strategy," *Acta Oto-Laryngol.* **129**, 988–991.
- Robles, L., Ruggero, M. A., and Rich, N. C. (1986). "Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases," *J. Acoust. Soc. Am.* **80**, 1364–1374.
- Rubinstein, J. T. (2004). "How cochlear implants encode speech," *Curr. Opin. Otolaryn.* **12**, 444–448.
- Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2012). "Why middle-aged listeners have trouble hearing in everyday settings," *Curr. Biol.* **22**, 1417–1422.
- Strelcyk, O., and Dau, T. (2009). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," *J. Acoust. Soc. Am.* **125**, 3328–3345.
- Swaminathan, J., and Heinz, M. G. (2012). "Psychophysiological analyses demonstrate the importance of neural envelope coding for speech perception in noise," *J. Neurosci.* **32**, 1747–1756.
- Vale, C., and Sanes, D. H. (2002). "The effect of bilateral deafness on excitatory and inhibitory synaptic strength in the inferior colliculus," *Eur. J. Neurosci.* **16**, 2394–2404.
- Wienbruch, C., Paul, I., Weisz, N., Elbert, T., and Roberts, L. E. (2006). "Frequency organization of the 40-Hz auditory steady-state response in normal hearing and in tinnitus," *Neuroimage* **33**, 180–194.
- Willott, J. F. (1991). "Central physiological correlates of ageing and presbycusis in mice," *Acta Oto-laryngol.* **111**, 153–156.