JARO
Journal of the Association for Research in Otolaryngology

Check for updates

Research Article

Age-Related Compensation Mechanism Revealed in the Cortical Representation of Degraded Speech

SAMIRA ANDERSON,[1] LINDSEY ROQUE,[1] CASEY R. GASKINS,[1] SANDRA GORDON-SALANT,[1] AND MATTHEW J. GOUPELL[1]

[1]*Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742, USA*

ABSTRACT

Older adults understand speech with comparative ease in quiet, but signal degradation can hinder speech understanding much more than it does in younger adults. This difficulty may result, in part, from temporal processing deficits related to the aging process and/or high-frequency hearing loss that can occur in listeners who have normal- or near-normal-hearing thresholds in the speech frequency range. Temporal processing deficits may manifest as degraded neural representation in peripheral and brainstem/midbrain structures that lead to compensation, or changes in response strength in auditory cortex. Little is understood about the process by which the neural representation of signals is improved or restored by age-related cortical compensation mechanisms. Therefore, we used vocoding to simulate spectral degradation to compare the behavioral and neural representation of words that contrast on a temporal dimension. Specifically, we used the closure duration of the silent interval between the vowel and the final affricate /tʃ/ or fricative /ʃ/ of the words DITCH and DISH, respectively. We obtained perceptual identification functions and electrophysiological neural measures (frequency-following responses (FFR) and cortical auditory-evoked potentials (CAEPs)) to unprocessed and vocoded versions of these words in young normal-hearing (YNH), older normal- or near-normal-hearing (ONH), and older hearing-impaired (OHI) listeners. We found that vocoding significantly reduced the slope of the perceptual identification function in only the OHI

listeners. In contrast to the limited effects of vocoding on perceptual performance, vocoding had robust effects on the FFRs across age groups, such that stimulus-to-response correlations and envelope magnitudes were significantly lower for vocoded vs. unprocessed conditions. Increases in the P1 peak amplitude for vocoded stimuli were found for both ONH and OHI listeners, but not for the YNH listeners. These results suggest that while vocoding substantially degrades early neural representation of speech stimuli in the midbrain, there may be cortical compensation in older listeners that is not seen in younger listeners.

Keywords: vocoding, midbrain, cortical, perception, aging

INTRODUCTION

Under ideal listening conditions, older adult listeners have little difficulty understanding conversational speech. The highly redundant nature of speech permits one to follow a conversation even when aspects of the speech signal are inaudible due to sensorineural hearing loss. When this redundancy is reduced by signal degradation, however, older listeners experience more difficulty understanding speech compared to younger listeners. This difficulty arises in part from decreased frequency resolution associated with hearing loss (Florentine et al. 1980; Phillips et al. 2000) and decreased temporal resolution associated with aging (Gordon-Salant and Fitzgibbons 1993; Gordon-Salant et al. 2007; Pichora-Fuller et al. 2007). The brain adapts to reduced or

Correspondence to: Samira Anderson · Department of Hearing and Speech Sciences · University of Maryland · College Park, MD 20742, USA. email: sander22@umd.edu

degraded input through homeostatic mechanisms; for example, changes in the balance of excitatory and inhibitory transmission occur in animals models of aging (Caspary et al. 2005; Hughes et al. 2010; Parthasarathy and Bartlett 2011) or hearing loss (Kotak et al. 2005; Dong et al. 2009). The effects of these compensatory neural mechanisms on perceptual and neural representations of speech are not well understood. Therefore, cortical compensation to degraded stimuli was investigated by comparing effects of vocoding on perceptual and neural representations in young normal-hearing (YNH), older normal- or near-normal-hearing (ONH), and older hearing-impaired (OHI) listeners.

Spectral degradation of speech signals, as occurs through vocoding, should increase the reliance on the temporal envelope to understand speech and thus highlight age-related changes in temporal processing abilities (Goupell et al. 2017). A compromised ability to distinguish between phoneme contrasts that differ based on a temporal cue is one example of how temporal processing is reduced by aging and hearing loss. Compared to younger listeners, older listeners exhibit poorer processing of temporal duration cues to distinguish phonemes that differ on a single acoustic cue, such as silence duration to cue the final affricate/fricative distinction in DITCH vs. DISH or vowel duration to cue the final voicing distinction in WHEAT vs. WEED (Gordon-Salant et al. 2006; Gordon-Salant et al. 2008; Goupell et al. 2017; Roque et al. 2019). The impaired ability of older listeners to utilize temporal (duration) cues is expected to reduce redundancy of speech cues and to negatively affect speech understanding (Gordon-Salant et al. 2011).

Effects of Spectral Degradation on Speech Perception

The effects of aging on the perception of spectrally degraded phonemes and words have previously been examined by using vocoder processing (Schvartz et al. 2008; Sheldon et al. 2008). Age was a primary factor for predicting recognition of vocoded vowels and stop consonants that had been spectrally shifted to mimic the frequency-to-place mismatch that occurs in cochlear-implant users (Schvartz et al. 2008). On a word recognition task, ONH listeners demonstrated reduced ability to use envelope cues compared to YNH listeners, but only for a task that involved single rather than multiple presentations of the word (Sheldon et al. 2008).

Goupell et al. (2017) also found that ONH listeners demonstrated reduced ability to discriminate words based on temporal envelope information compared to YNH listeners. They obtained perceptual identification functions in YNH and ONH listeners for a continuum of speech tokens that differed in the silence duration preceding the final fricative in 10-ms increments (0-ms interval perceived as DISH and 60-ms interval perceived as DITCH). These tokens were presented in unprocessed and vocoded conditions. The stimuli were vocoded using a sine-wave carrier; the stimuli varied in the number of channels (1, 2, 4, 8, and 16) and temporal envelope low-pass cut-off frequency (50 and 400 Hz). They found that ONH listeners required a longer silent interval to perceive the token as DITCH than YNH listeners and that these age-related differences became particularly large as spectral information was reduced to 8 or 4 channels.

Effects of Signal Degradation on Neural Representation of Speech

Effects of spectral degradation on neural responses have been evaluated using different electrophysiological approaches. Ananthakrishnan et al. (2017) recorded the frequency-following response (FFR) to a vocoded vowel to evaluate the effects of spectral degradation on subcortical representation of the fundamental frequency (F_0) and the first speech formant (F_1) in YNH listeners. The synthesized vowel /u/ in that study was processed using vocoders that varied in the carrier (sine or noise band), number of channels (1, 2, 4, 8, 16, or 32), and temporal envelope cut-off frequency (50 or 500 Hz). They found that the magnitude of the F_0 and the F_1 in the FFR increased as the number of channels increased from 1 to 4. From 8 to 16 channels, the magnitude plateaued and then decreased with 32 channels. These results suggest that neural phase locking to the F_0 improves with the greater spectral resolution afforded by increasing channels; the decreased magnitude for 32 channels likely results from narrowing the channel bandwidths, which would limit the envelope modulation frequencies passed through the band-pass filters.

Friesen et al. (2009) used cortical auditory-evoked potentials (CAEPs) to evaluate spectral degradation effects on cortical processing of consonant-vowel-consonant (CVC) stimuli in YNH listeners. The CVC stimuli were noise vocoded with 2, 4, 8, 12, and 16 channels. They found that vocoding affected neural representation of speech stimuli; specifically, the cortical peaks increased in amplitude and decreased in latency as the number of channels increased.

The Ananthakrishnan et al. (2017) and Friesen et al. (2009) studies investigated effects of spectral degradation in YNH listeners on subcortical and cortical responses, respectively. However, ONH and OHI listeners may differ in the degree to which their responses are affected by stimulus degradation at subcortical and cortical levels compared to YNH

listeners. At the subcortical level, older listeners have reduced FFR amplitudes compared to younger listeners. In addition, competing noise, which typically decreases amplitudes in younger listeners, does not decrease amplitudes to the same degree in ONH compared to YNH listeners (Presacco et al. 2016b). In contrast, ONH and OHI listeners' neural responses in auditory cortex show exaggerated or over-representation of speech stimuli (higher reconstruction accuracy) compared to YNH listeners, but noise has similar effects (decreased reconstruction accuracy) across age groups (Brodbeck et al. 2018; Presacco et al. 2019).

Assuming that spectral degradation affects neural responses in a manner similar to the Presacco et al. (2019) study, it was hypothesized that vocoding results in greater reductions in morphology/phase locking in subcortical responses in YNH compared to ONH and OHI listeners. We also hypothesized that vocoding would result in exaggerated amplitudes in ONH and OHI listeners, but not in YNH listeners. To test these hypotheses, FFRs and CAEPs were recorded to vocoded and unprocessed words that differed on a silence duration cue (DISH vs. DITCH) in YNH, ONH, and OHI listeners. Perceptual identification functions were obtained to vocoded and unprocessed DISH-DITCH tokens on a continuum of silent interval durations, and linear regressions were used to determine the relationship between neural representation and behavioral performance.

METHODS

Listeners

Three groups of native English speakers were recruited from the Washington, D.C., Maryland, and Virginia areas including 15 YNH (9 female, 18–26 years, average = 21.27 years, standard deviation = 2.22 years), 15 older listeners with normal- or near-normal-hearing thresholds (ONH; 12 female, 61–78 years, average 68.32 years, standard deviation 4.75 years), and 15 OHI (8 females, 63–81 years, average = 73.86 years, standard deviation = 5.10 years) listeners. There were no significant differences in sex between groups ($\chi^2 = 2.47$, $P = 0.29$). Average audiometric thresholds for each listener group are shown in Fig. 1. Audiometrically normal- or near-normal-hearing was defined as hearing thresholds ≤ 25 dB HL at octave frequencies from 125 to 4000 Hz and interaural asymmetries < 15 dB HL at two or fewer adjacent frequencies. OHI listeners demonstrated pure-tone averages ≥ 25 dB HL for octave frequencies from 500 to 4000 Hz. All listeners were screened for mild cognitive dysfunction: the inclusion criterion was a score of ≥ 22 out of a possible 30 points on the Montreal Cognitive Assessment (MoCA Version 7.1 Original
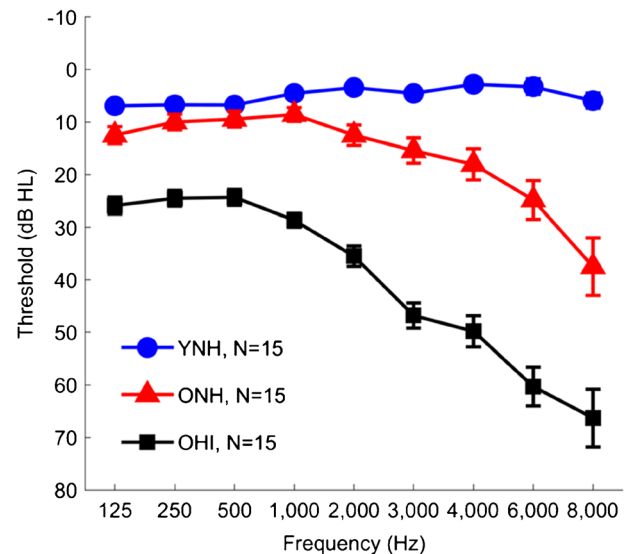
Fig. 1. Group average thresholds in young normal-hearing (YNH, blue circles), older normal-hearing (ONH, red triangles), and older hearing-impaired (OHI, black squares) listeners. Although the ONH listeners have clinically normal hearing, their thresholds are elevated compared to the YNH listeners across the frequency range. Thresholds in the OHI listeners range from mild hearing loss levels in the low frequencies sloping to moderate to severe hearing loss levels in the high frequencies. Error bars show ± 1 standard error. The number of subjects for each group are displayed with the legend

Version; Nasreddine et al. 2005). YNH, ONH, and OHI listeners obtained mean scores and standard deviations of 29.3 ± 1.0, 26.9 ± 2.5, and 27.2 ± 1.7, respectively. A significant main effect of age was observed on MoCA score ($F_{(2,44)} = 5.24$, $P = 0.009$). Follow-up independent samples t tests showed higher scores in the YNH compared to the ONH ($P = 0.018$) and OHI ($P = 0.042$) listeners. No difference in MoCA score was observed between the ONH and OHI listeners ($P = 0.93$). Listeners also had a normal intelligence quotient (≥ 85) on the Wechsler Abbreviated Scale of Intelligence (WASI; Zhu and Garcia 1999). Mean scores and standard deviations for YNH, ONH, and OHI listeners were 115.5 ± 9.6, 120.5 ± 15.4, and 111.1 ± 13.0, respectively, with no significant difference between groups ($F_{(2,44)} = 1.61$, $P = 0.21$). Listeners with a history of neurological dysfunction or middle ear surgery were excluded from the study. Auditory brainstem responses were recorded to 100-μs clicks presented at 80 dB peSPL at 21.1 Hz using the Intelligent Hearing System SmartEP system (IHS, Miami, FL) to verify auditory neural integrity. All listeners had wave V latencies ≤ 6.5 ms and interaural differences < 0.2 ms. The procedures were reviewed and approved by the Institutional Review Board of the University of Maryland, College Park. Listeners gave informed consent and were compensated for their time.
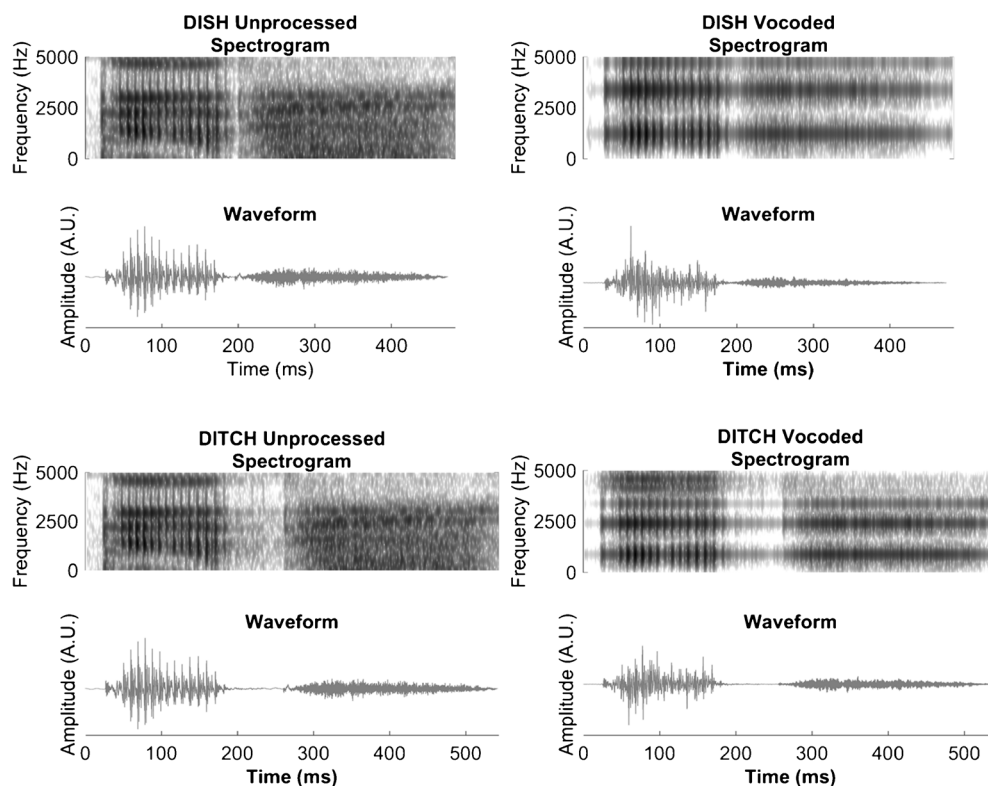
Stimuli

Test stimuli were speech tokens of the contrasting word pair DISH (483 ms) and DITCH (543 ms), first described in Gordon-Salant et al. (2006). The DISH-DITCH contrast was chosen because the silence duration parameter yielded the largest group differences in a study that examined aging and hearing loss effects on behavioral performance for four duration contrasts: silence duration, vowel duration, consonant transition duration, and voice onset time (Gordon-Salant et al. 2006). The speech tokens were generated from isolated recordings of the natural words DISH and DITCH as spoken by an adult American male. The DISH-DITCH word pair was chosen because it varies in the duration of the silent interval preceding the final fricative (0 ms) in DISH or affricate (60 ms) in DITCH. The endpoint of the continuum perceived as DITCH was a hybrid word, in which the burst and final affricate /tʃ/ were excised and replaced with the final fricative /ʃ/ from the natural speech token DISH. The resulting DITCH endpoint comprised an initial stop, vowel, closure, and final fricative. A seven-step continuum was then created by excising 10-ms intervals of silence from the closure interval in DITCH until reaching 0 ms of silence (the DISH endpoint).

Once the DISH-DITCH continuum was created, it was spectrally degraded using vocoding. The vocoded DISH-DITCH continuum was first described in Goupell et al. (2017). Each step of the unprocessed continuum was band-pass filtered into eight contiguous channels by the vocoder using third-order Butterworth filters. The corner frequencies of the band-pass filters were logarithmically spaced from 200 to 8000 Hz. To better preserve the temporal acoustic cues in speech, forward-backward filtering was performed, doubling the filter order and resulting in a −36 dB/octave attenuation slope. The envelopes of the stimuli were extracted by calculating the Hilbert envelope and low-pass forward-backward filtering using a second-order Butterworth filter with a 400-Hz envelope cut-off. The extracted envelopes were used to modulate the sinusoidal carriers at the geometric mean frequency of each band-pass filter and were then summed across channels. The final vocoded stimuli were normalized so that they had the same root-mean-square amplitude as the unprocessed stimuli. Figure 2 displays the spectrograms and waveforms for the unprocessed and vocoded DISH and DITCH stimuli.

For the perceptual identification functions, listeners were presented with both the unprocessed and vocoded DISH-DITCH continua of silence duration, ranging from 0 ms (DISH) to 60 ms (DITCH), preceding the final fricative /ʃ/. For the electrophys-

FIG. 2. Unprocessed and vocoded spectrograms and stimulus waveforms for DISH and DITCH. A reduction in periodicity is noted in vocoded waveforms

iology recordings, listeners were only presented with the two tokens that were the endpoints of the DISH-DITCH continua.

Procedure

Perceptual. Unprocessed and vocoded DISH-DITCH continua were presented to listeners in an identification task similar to that utilized in Gordon-Salant et al. (2006). The experiment was implemented, and listener responses were recorded using MATLAB (MathWorks, version 2012a). During the experiment, the listener was seated at a desktop computer in a double-walled sound-attenuated booth (IAC Acoustics, North Aurora, IL). Three boxes were displayed on the computer monitor: one that read "Begin Trial" and two boxes below it that read "DISH" and "DITCH." The listener initiated each trial by clicking the "Begin Trial" box; testing was therefore self-paced. Stimuli were presented monaurally to the right ear via an ER-2 insert earphone (Etymotic Research, Elk Grove Village, IL) at 75 dB SPL. Following each presentation, the listener indicated whether the stimulus was perceived as DISH or DITCH by clicking on the corresponding box. Listeners were encouraged to guess if they were unsure of which stimulus was presented and to distribute guesses equally between DISH and DITCH. Before experimental testing, each listener completed a training run that presented the continuum endpoints in the unprocessed condition and provided feedback after each trial. Listeners had to achieve 90 % accuracy before initiating the experimental runs. Two of the OHI listeners were unable to achieve this level of accuracy; their data were not included in the perceptual analysis, but they were included in the electrophysiological analyses because clear response waveforms were obtained in both these listeners. During the experimental testing, the listeners completed five blocks and were not provided with feedback. Stimuli along both DISH-DITCH continua (unprocessed and vocoded) were each presented a total of ten times over the course of the experimental blocks.

Electrophysiological

Subcortical. Presentation software (Neurobehavioral Systems, Berkeley, CA) was used to present the endpoints of the unprocessed and vocoded DISH-DITCH continua to each listener. Stimuli were presented monaurally to the right ear via an ER-1 insert earphone (Etymotic Research, Elk Grove Village, IL) at 75 dB SPL. The 28-ms final fricative /ʃ/ in DISH was excised in Adobe Audition (Adobe, San Jose, CA) and saved as a separate waveform. A fast Fourier transform was calculated on the waveform and the dominant energy was > 2 kHz. Prior to the recording, thresholds for each listener were obtained for the excised /ʃ/ stimulus to ensure that this high-frequency token was sufficiently audible. Listeners were only included if they had thresholds of ≤ 55 dB SPL for the /ʃ/ token, thus ensuring that the DISH-DITCH tokens were highly audible at a presentation level that was at least 20 dB SL re: the detection threshold for the fricative. For each speech token, the entire stimulus waveform was presented with alternating polarities at a rate of 1.5 Hz. The BioSemi ActiABR-200 acquisition system (BioSemi B.V., Netherlands) was used to record FFRs. Recordings were obtained with a standard five-electrode vertical montage (Cz active, two forehead offset CMS/DRL electrodes, two earlobe reference electrodes) at a sampling rate of 16,384 Hz.

Cortical. The endpoints of the DISH-DITCH continua also served as the stimuli for the CAEP recordings. Stimuli were presented at a rate of 0.71 Hz monaurally to the right ear at 75 dB SPL. Recordings were obtained with a 32-channel electrode cap with average earlobe electrodes (A1 and A2) serving as references using the BioSemi ActiveTwo system. A minimum of 500 artifact-free sweeps were recorded at a sampling rate of 2048 Hz for each stimulus from each listener.

Data Analysis

Perceptual. Percent identification of DISH responses was calculated for each step of the unprocessed and vocoded continua. The PSIGNIFIT software (https://sourceforge.net/projects/psignifit/) was used to calculate the crossover point and the slope over the entire function for each listener, using the same procedure described in Wichmann and Hill (2001).

Electrophysiological

FFR Data Reduction. All electrophysiological (EEG) data analyses were completed in MATLAB (MathWorks, version R2011b). Recordings were first converted into MATLAB format using the pop_biosig function from EEGLAB (Delorme and Makeig 2004). Sweeps within the 30-μV amplitude range were retained. Recordings were band-pass filtered offline from 70 to 2000 Hz using a zero-phase, fourth-order Butterworth filter. To maximize the temporal envelope of the response waveforms, a final average response was created by averaging the sweeps of both polarities over a 600-ms time window.

Stimulus-to-Response (STR) Correlations. STR correlations (r values) were obtained to examine the degree to which listener response waveform envelopes approximated those of the stimulus waveform envelopes. The envelope of the stimulus waveform was extracted and band-pass filtered with the same parameters as were applied to the response waveforms (70–2000 Hz). The

XCorr function in MATLAB was used to obtain the STR correlation by shifting the stimulus waveform in time from 10 to 300 ms relative to the response waveform until a maximum correlation was found.

Fast Fourier Transform (FFT). FFTs were used to decompose FFRs into their individual frequency components to examine phase locking to the temporal envelope. Average spectral magnitudes over 20-Hz bins were calculated from individual responses with zero padding and 1-Hz interpolated frequency resolution over the vowel region prior to the silent interval (50–190 ms) for the 110-Hz fundamental frequency (F_0).

CAEP. Recordings were band-pass filtered offline from 1 to 30 Hz using a zero-phase, fourth order Butterworth filter. A regression-based electrooculography reduction method was implemented to remove eye movements from the filtered data (Romero et al. 2006; Schlögl et al. 2007). A time window of – 500 to 1000 ms was referenced to the stimulus onset for each sweep. A final response was created from an average of the first 500 artifact-free sweeps. The denoising source separation algorithm was utilized to minimize artifact in data obtained from the 32 channels recorded (Särelä and Valpola 2005; de Cheveigne and Simon 2008). An automated algorithm was utilized in MATLAB to identify the latencies and amplitudes of prominent CAEP peaks in the expected time region based on the average response for YNH, ONH, and OHI listeners. The expected time regions were as follows: P1 (40–90 ms), N1 (100–150 ms), P2 (150–210 ms), and P1b (210–270 ms) for YNH listeners and P1 (50–100 ms), N1 (100–150 ms), P2 (180–270 ms), and P1b (290–340 ms) for ONH and OHI listeners. Note that the P1b component corresponds to the onset of the fricative /ʃ/.

Statistical Analysis

All statistical analyses were completed in JASP (JASP Team 2018). Split-plot analyses of variance (ANOVA) were conducted to assess between-group differences (YNH vs. ONH vs. OHI) and within-group effects of vocoding (unprocessed vs. vocoded) for perceptual crossover point and slope. In addition, split-plot ANOVAs were completed for evaluating between-group comparisons (YNH vs. ONH vs. OHI), and within-group effects of stimulus (DISH vs. DITCH) and vocoding (unprocessed vs. vocoded) on the subcortical FFR variables (STR and FFT) in the time region preceding the silent interval (50–190 ms). A similar analysis was performed on the CAEP variables (peak latency and amplitude of P1, N1, P2, and P1b). Sheffé's post hoc analyses were performed when

significant main effects or interactions were observed. For STR correlations, Fisher's z-transformation was applied prior to statistical analysis.

Pearson's correlations were performed to examine relationships among the perceptual, peripheral, subcortical, and cortical measures. Multiple comparisons were addressed with the false discovery rate (Benjamini and Hochberg 1995). Separate multiple linear regressions were performed on the vocoded 50 % crossover point for the DISH-DITCH continuum serving as the dependent variable. The pure-tone average (PTA, average of right ear thresholds at 500, 1000, 2000, 3000, and 4000 Hz) and difference measures (vocoded minus unprocessed) for STR correlation and P1 amplitude served as the independent variables. Note that the PTA measure was not normally distributed, so the log-transform of this measure was used. The appropriateness of the linear regression analysis for the data set was verified by checking the residuals for normality. Collinearity diagnostics indicated a satisfactory variance inflation factor (highest = 1.08), denoting the absence of strong correlations between two or more predictors.

RESULTS

Perceptual

Average unprocessed and vocoded identification functions for YNH, ONH, and OHI listeners are shown in Fig. 3. Individual 50 % crossover points and slopes were calculated for perceptual judgments along each continuum of silence duration and averaged for each group. Two ONH listeners and one OHI listener had nearly flat identification functions for the vocoded continuum, and their crossover points were set at 60 ms. Figure 4 displays perceptual identification functions (top row) and individual and mean average values for the 50 % crossover points (middle row) and slopes (bottom row), contrasting unprocessed and vocoded stimuli for each listener group. Mean 50 % crossover points and slopes, along with standard deviations, for the three listener groups to the unprocessed and vocoded continua are shown in Table 1. A split-plot ANOVA revealed no main effect of vocoding on the crossover point ($F_{(1,40)}$ = 1.37, P = 0.25, η^2 = 0.03). However, a main effect of group was observed for the crossover point ($F_{(2,40)}$ = 14.37, P < 0.001, η^2 = 0.42). Sheffe's post hoc analyses revealed that YNH listeners exhibited earlier 50 % crossover points relative to both ONH (P = 0.018) and OHI listeners (P < 0.001), with no differences observed between the two groups of older listeners (P = 0.43). There was no significant group × vocoding interaction ($F_{(2,40)}$ = 0.26, P = 0.85, η^2 = 0.01).
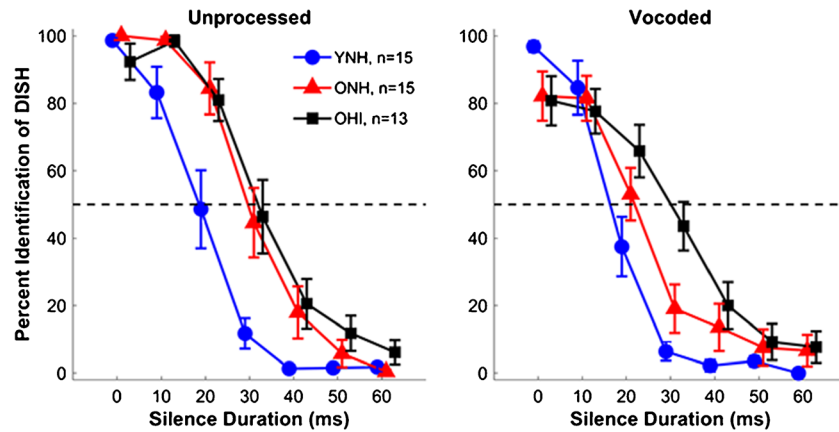
FIG. 3. Split-plot ANOVA revealed aging effects: YNH listeners show earlier 50 % crossover points on perceptual identification functions than ONH and OHI listeners ($P < 0.001$), but no differences are noted between ONH and OHI listeners. YNH listeners also show steeper slopes than the OHI listeners ($P = 0.005$), but no other groups differences were noted (all P values > 0.05). Group average comparisons for perceptual identification functions are displayed for unprocessed and vocoded stimuli obtained from YNH (blue), ONH (red), OHI (black) listeners. Error bars show ± 1 standard error. The Ns for each group are displayed with the legend in the left panel. Note that two OHI listeners were not included in the perceptual data as they were unable to complete the training task
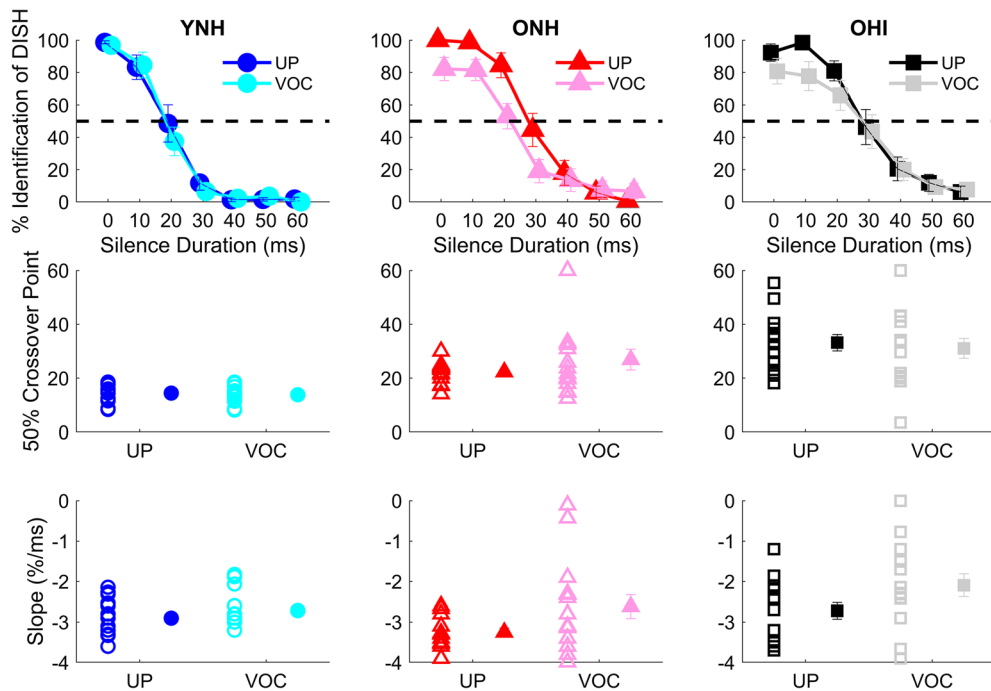
FIG. 4. Split-plot ANOVA revealed minimal effects of vocoding: ONH listeners had shallower slopes in the vocoded compared to the unprocessed conditions ($P = 0.001$), but no other vocoding effects were found for slope or crossover point. Top row: group average perceptual identification functions in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners for unprocessed and vocoded stimuli. Middle row: individual (open symbols) and group mean (closed symbols) values for 50 % crossover points for the perceptual identification functions. Bottom row: individual (open symbols) and group mean (closed symbols) values for slope to vocoded and unprocessed DISH and DITCH stimuli. Symbols: YNH: unprocessed = blue circles, vocoded = cyan circles; ONH: unprocessed = red triangles, vocoded = pink triangles; OHI: unprocessed = black squares, vocoded = gray squares. Error bars show ± 1 standard error. **$P < 0.01$

Group	50 % crossover point		Slope	
	Unprocessed	Vocoded	Unprocessed	Vocoded
YNH	19.0 (6.6)	17.8 (5.8)	-2.90 (0.13)	-2.72 (0.11)
ONH	29.8 (8.6)	25.8 (16.6)	-3.25 (0.10)	-2.62 (0.29)
OHI	33.2 (11.0)	31.0 (14.3)	-2.72 (0.21)	-2.09 (0.28)

There was a main effect of vocoding on slope ($F_{(1,40)} = 10.25$, $P < 0.001$, $\eta^2 = 0.21$) and a main effect of group ($F_{(2,40)} = 6.20$, $P = 0.005$, $\eta^2 = 0.24$). YNH listeners demonstrated steeper slopes compared to OHI listeners ($P = 0.005$), but there was neither a group difference between YNH and ONH listeners nor between ONH and OHI listeners (all P values > 0.05). The group × vocoding interaction was significant ($F_{(2,40)} = 4.29$, $P = 0.02$, $\eta^2 = 0.14$), driven by an effect of vocoding on slope that was seen in the ONH listeners ($P = 0.001$) but not seen in the YNH or OHI listeners (both P values > 0.05). As can be observed in the top panel of Fig. 4, the performance of the ONH listeners did not reach 100 % for DISH nor 0 % for DITCH at the endpoints of the vocoded continuum, explaining the relatively shallow slope. The YNH listeners did not show a difference in performance between the vocoded and unprocessed endpoint tokens, while the OHI listeners exhibited a slight difference in recognition between vocoded and unprocessed DISH but not DITCH.

Upon further examination of the identification functions in Fig. 4, we noted that the slopes in the linear portions of the functions appeared to be equivalent between unprocessed and vocoded conditions in the ONH listeners. The PSIGNIFIT software calculates slope from the entire function, including the endpoints, and because performance in the ONH listeners did not reach 100 % for DISH or 0 % for DITCH, the slope differences were likely due to vocoding-related changes in recognition of DISH-DITCH at or near the endpoints. For this reason, we recalculated the slope by performing a linear regression using data points only from the linear portion of each identification function. The results of this calculation are displayed in Fig. 4, bottom panel. Using this method of slope calculation, there was a main effect of vocoding ($F_{(1,40)} = 10.57$, $P = 0.002$, $\eta^2 = 0.09$) but no main effect of group ($F_{(2,40)} = 2.83$, $P = 0.07$, $\eta^2 = 0.12$). Although the group × vocoding interaction was not significant ($F_{(2,40)} = 1.03$, $P = 0.37$, $\eta^2 = 0.02$), we performed planned paired t tests to determine if the vocoding effects for each group were

observed with this analysis. We found that vocoding effects were significant for the OHI listeners ($P = 0.04$) but not for the YNH ($P = 0.28$) or ONH ($P = 0.06$) listeners. Thus, a slope analysis confined to the linear portion of the function indicates an effect of vocoding primarily for the OHI listeners, although the slope analysis that considers all points in the function indicates an effect of vocoding primarily for the ONH listeners. While these results may appear to be inconsistent, they generally suggest that listeners with hearing loss are more susceptible to the effects of vocoding, making distinctions of speech stimuli based on a temporal cue somewhat obscured for these listeners, whereas vocoding does not affect the ability to use temporal cues in listeners with normal hearing.

## Electrophysiology

Subcortical

FFR STR Correlation. Average response waveforms of the three listener groups to unprocessed and vocoded stimuli are shown in Fig. 5. The data are quantified in Fig. 6. A main effect of vocoding was found on STR ($F_{(1,42)} = 109.49$, $P < 0.001$, $\eta^2 = 0.71$), such that there was better waveform morphology to unprocessed stimuli compared to vocoded stimuli (Fig. 5), quantified by higher r values (Fig. 6). There was a main effect of group ($F_{(2,42)} = 3.32$, $P = 0.046$, $\eta^2 = 0.14$) on the STR correlations (Fig. 6). Follow-up t tests revealed higher morphology for the unprocessed DITCH stimulus in YNH vs. ONH listeners ($P = 0.008$) and in YNH vs. OHI listeners ($P = 0.047$). No significant group differences were observed for the vocoded DITCH stimulus or for the unprocessed and vocoded DISH stimuli (all P values > 0.05) (Fig. 6). There was no significant vocoding × group interaction ($F_{(2,42)} = 1.68$, $P = 0.20$, $\eta^2 = 0.02$) or word × vocoding × group interaction ($F_{(2,42)} = 1.78$, $P = 0.18$, $\eta^2 = 0.01$).

FFT. Figure 7 shows average temporal envelope spectral magnitudes for the three listener groups. A main effect of vocoding was observed on the F_0 ($F_{(1,42)} = 35.63$, $P < 0.001$, $\eta^2 = 0.71$), such that spectral magnitudes for unprocessed stimulus envelopes were greater than for vocoded stimulus envelopes. Neither the main effect of group nor the interaction between vocoding and group was significant (all P values > 0.05).

Cortical

Amplitude. Figure 8 displays average CAEP response waveforms for YNH, ONH, and OHI listeners. Figure 9 displays the quantified CAEP amplitudes for prominent peak components for the same listeners. There was no main effect of vocoding observed on CAEP amplitude ($F_{(1,37)} = 0.09 = 0.76$, $\eta^2 = 0.02$). However, there was a significant peak × vocoding interac-
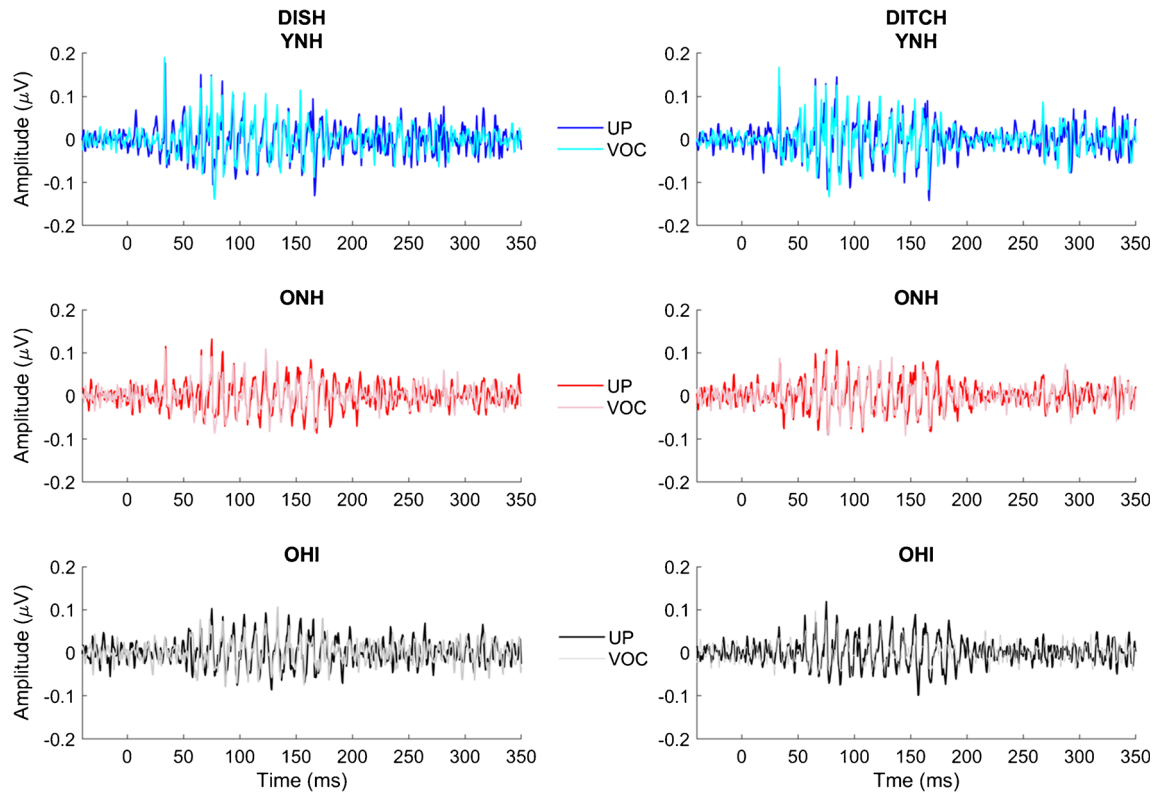
FIG. 5. Response waveforms mimic stimulus waveforms (displayed in Fig. 2) across groups and conditions. Group average response waveforms to DISH and DITCH are displayed in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners, YNH: unprocessed = blue, vocoded = cyan; ONH: unprocessed = red, vocoded = pink; OHI: unprocessed = black, vocoded = gray

tion ($F_{(2,37)} = 15.12$, $P = <0.001$, $\eta^2 = 0.27$). To further understand this interaction, we conducted follow-up split-plot ANOVAs for the four individual peaks, with vocoding and stimulus as two within-subject factors.

P1. There was a main effect of vocoding ($F_{(1,38)} = 33.93$, $P < 0.001$, $\eta^2 = 0.41$), due to larger amplitudes in the vocoded compared to the unprocessed conditions. There was a main effect of group ($F_{(2,38)} = 5.90$, $P < 0.01$, $\eta^2 = 0.24$). Post hoc t tests indicated larger amplitudes in the ONH listeners compared to the YNH listeners ($P < 0.01$), but no significant differences between the ONH and OHI listeners ($P = 0.48$) nor between the YNH and OHI listeners ($P = 0.19$). There was a significant vocoding × group interaction ($F_{(2,38)} = 5.34$, $P < 0.01$, $\eta^2 = 0.13$). This interaction was driven by a vocoding-related increase in P1 amplitude in the ONH listeners ($F_{(1,13)} = 35.55$, $P < 0.001$, $\eta^2 = 0.73$) and the OHI listeners ($F_{(1,13)} = 10.94$, $P = 0.006$, $\eta^2 = 0.48$) that was not seen in the YNH listeners ($F_{(1,13)} = 0.85$, $P = 0.37$, $\eta^2 = 0.06$).

N1. There was a main effect of vocoding ($F_{(1,38)} = 4.61$, $P = 0.038$, $\eta^2 = 0.11$), driven by a decrease in amplitude in the vocoded compared to the unprocessed condition across groups, and there was a significant effect of group, such that the ONH and OHI listeners had larger N1 amplitudes than the YNH listeners ($F_{(2,38)} = 27.94$, $P < 0.001$, $\eta^2 = 0.60$). Post hoc testing revealed that OHI listeners had larger amplitudes than YNH and ONH listeners (both P values < 0.001), but there was no significant amplitude difference between YNH and ONH listeners ($P = 0.07$). The interaction between vocoding and group was not significant ($F_{(2,38)} = 0.54$, $P = 0.59$, $\eta^2 = 0.03$).

P2. None of the main effects nor the interaction were significant (vocoding: $F_{(1,38)} = 2.18$, $P = 0.148$, $\eta^2 = 0.05$; group: $F_{(1,38)} = 036$, $P = 0.70$, $\eta^2 = 0.02$; vocoding × group interaction: $F_{(2,38)} = 1.25$, $P = 0.30$, $\eta^2 = 0.06$).

P1b. There was a main effect of vocoding ($F_{(1,38)} = 6.26$, $P = 0.017$, $\eta^2 = 0.14$), driven by a decrease in amplitude in the vocoded condition compared to the unprocessed condition across groups. There was also a significant main effect of group ($F_{(2,38)} = 12.90$, $P < 0.001$, $\eta^2 = 0.40$). Post hoc testing revealed that the YNH listeners had larger amplitudes than ONH
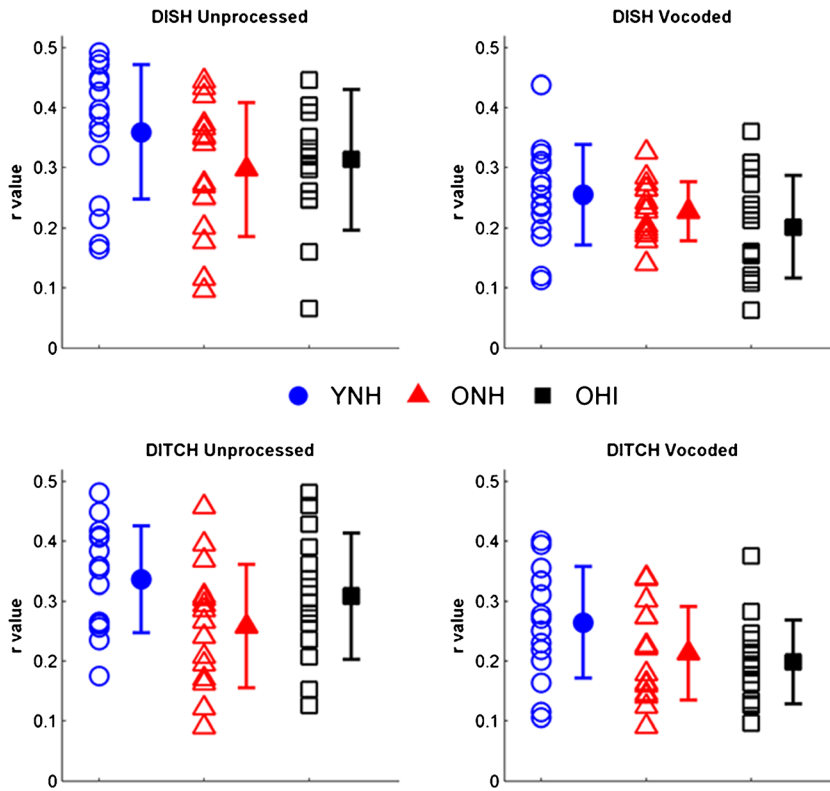
FIG. 6. Vocoding reductions in morphology differ across groups (*P*<0.001). Older listeners have lower morphology compared to the YNH listeners, but this effect is specific to the unprocessed DITCH condition (YNH vs. ONH, *P*=0.008; YNH vs. OHI, *P*=0.047). Individual (open symbols) and group mean (closed symbols) for stimulus-to-response correlation *r* values to DISH (top row) and DITCH (bottom row) in unprocessed (left column) and vocoded (right column) conditions in YNH (blue circles), ONH (red triangles), and OHI (black squares) listeners. Error bars show ±1 standard deviation
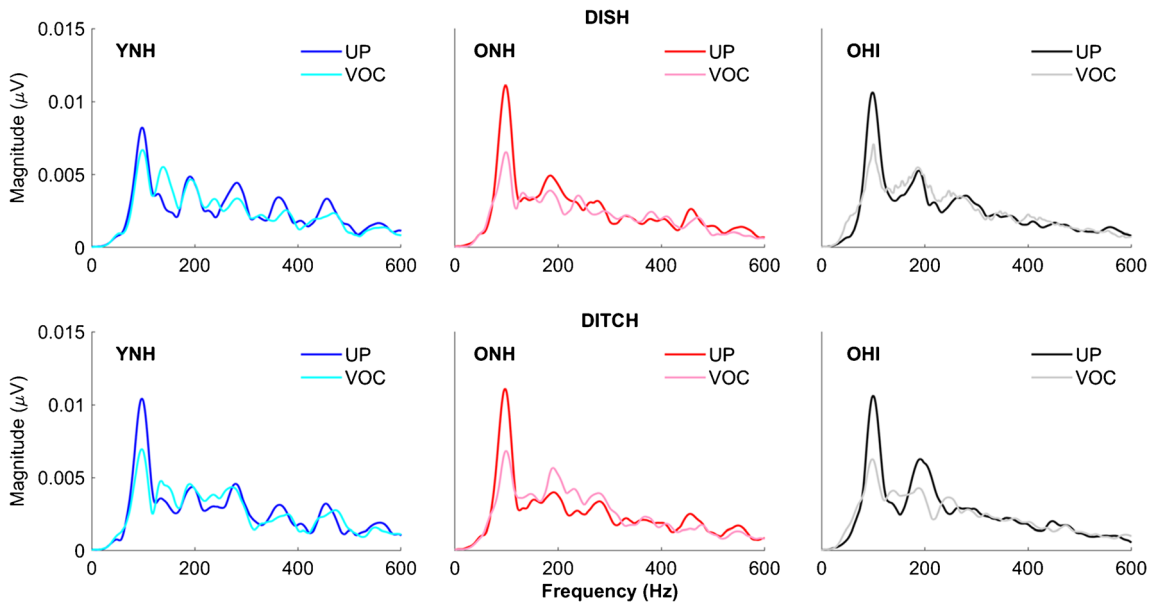

FIG. 7. Vocoding reduces fundamental frequency magnitude across groups (*P*<0.001). Group average spectral magnitudes for unprocessed and vocoded stimuli in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners for both DISH and DITCH stimuli. YNH: unprocessed = blue, vocoded = cyan; ONH: unprocessed = red, vocoded = pink; OHI: unprocessed = black squares, vocoded = gray squares
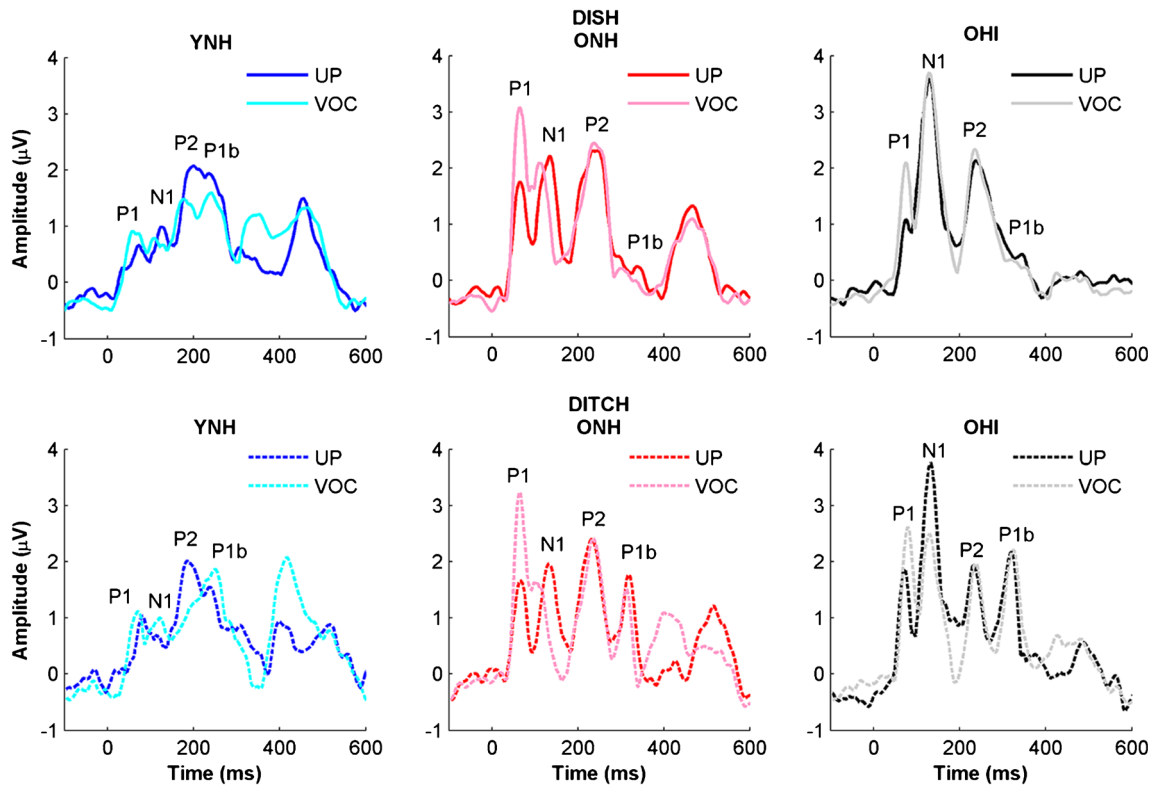
FIG. 8. Group average cortical waveforms obtained using the denoising source separation algorithm for unprocessed and vocoded DISH and DITCH stimuli in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners. Expected aging effects are observed for more prominent N1 and P2 components in the older listeners compared to the younger listeners. Note that the denoising source separation algorithm rectified the peaks, so the typically negative N1 component is a positive peak. YNH: unprocessed = blue, vocoded = cyan; ONH: unprocessed = red, vocoded = pink; OHI: unprocessed = black; vocoded = gray
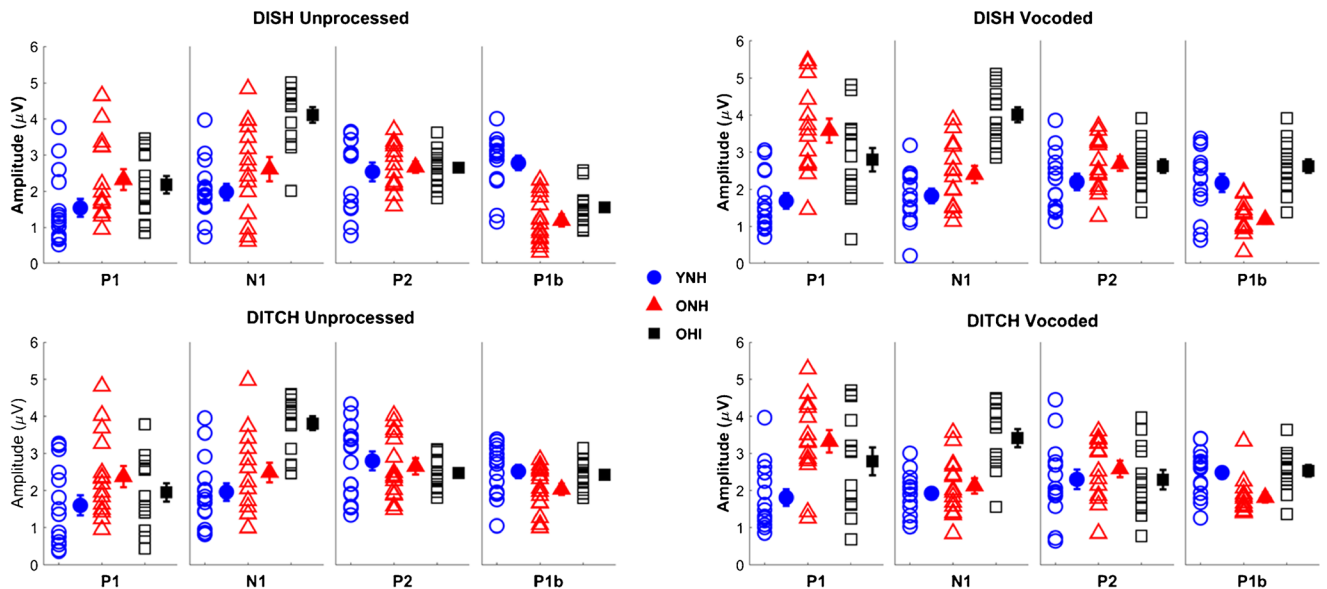
FIG. 9. Split-plot ANOVAs revealed that vocoding increases P1 amplitudes in both ONH ($P < 0.001$) and OHI ($P < 0.01$) listeners but not in YNH listeners ($P = 0.37$). In contrast, vocoding decreases N1 amplitudes across groups ($P = 0.038$). Individual (open symbols) and group mean (closed symbols) for cortical amplitudes to DISH (top panels) and DITCH (bottom panels) in unprocessed (left column) and vocoded (right column) conditions in YNH (blue circles), ONH (red triangles), and OHI (black squares) listeners. Error bars show ±1 standard error

listeners ($P < 0.001$) and OHI listeners ($P = 0.022$), but there was no difference in amplitude between ONH and OHI listeners ($P = 0.117$). There was no significant vocoding × group interaction ($F_{(2,38)} = 1.08$, $P = 0.35$, $\eta^2 = 0.05$).

Latency. Figure 10 displays average CAEP latencies for prominent peak components for YNH, ONH, and OHI listeners. There was no main effect of vocoding observed on CAEP latency ($F_{(1,38)} = 3.35$, $P = 0.075$, $\eta^2 = 0.08$). However, there was a peak × vocoding interaction ($F_{(3,36)} = 3.80$, $P = 0.049$, $\eta^2 = 0.06$). As a result, we conducted follow-up analyses for individual peaks, with vocoding and stimulus as two within-subject factors.

P1. There was no main effect of vocoding ($F_{(1,38)} = 2.78$, $P = 0.10$, $\eta^2 < 0.01$). There was a significant effect of group ($F_{(2,38)} = 3.89$, $P = 0.03$, $\eta^2 = 0.17$). Post hoc testing revealed that P1 latencies were earlier in the ONH than in the OHI listeners ($P = 0.048$), but no significant differences in latency were observed between the YNH and ONH listeners ($P = 1.00$) nor between the YNH and OHI listeners ($P = 0.08$). There was no significant vocoding × group interaction ($F_{(2,38)} = 0.14$, $P = 0.87$, $\eta^2 = 0.01$).

N1. There was a main effect of vocoding ($F_{(1,38)} = 10.43$, $P < 0.01$, $\eta^2 = 0.20$), driven by earlier latencies in the vocoded compared to the unprocessed condition across listeners. There was no significant effect of group ($F_{(2,38)} = 2.64$, $P = 0.09$, $\eta^2 = 0.12$) nor a

significant vocoding × group interaction ($F_{(2,38)} = 1.42$, $P = 0.23$, $\eta^2 = 0.06$).

P2. There was no main effect of vocoding ($F_{(1,38)} = 0.03$, $P = 0.86$, $\eta^2 < 0.01$). There was a significant effect of group, such that the younger listeners had earlier P2 latencies than both the ONH and OHI listeners ($F_{(2,38)} = 66.04$, $P < 0.001$, $\eta^2 = 0.78$), and post hoc testing revealed that this effect was significant between the YNH and ONH listeners ($P < 0.001$) and between the YNH and OHI listeners ($P < 0.001$), but there was no latency difference between the ONH and OHI listeners ($P = 1.00$). There was no significant vocoding × group interaction ($F_{(2,38)} = 0.66$, $P = 0.525$, $\eta^2 = 0.03$).

P1b. There was no main effect of vocoding ($F_{(1,38)} = 1.42$, $P = 0.24$, $\eta^2 = 0.03$). There was a significant effect of group, such that the younger listeners had earlier P1b latencies than both the ONH and OHI listeners ($F_{(2,38)} = 193.7$, $P < 0.001$, $\eta^2 = 0.91$), and post hoc testing revealed that this effect was significant between the YNH and ONH listeners ($P < 0.001$) and between the YNH and OHI listeners ($P < 0.001$), but there was no latency difference between the ONH and OHI listeners ($P = 0.97$). There was no significant vocoding × group interaction ($F_{(2,38)} = 1.25$, $P = 0.30$, $\eta^2 = 0.06$).

Cortical Summary. Vocoding effects on amplitude differed for specific peaks. P1 amplitudes increased with vocoding, but only in the two groups of older listeners. N1 and P1b amplitudes decreased with
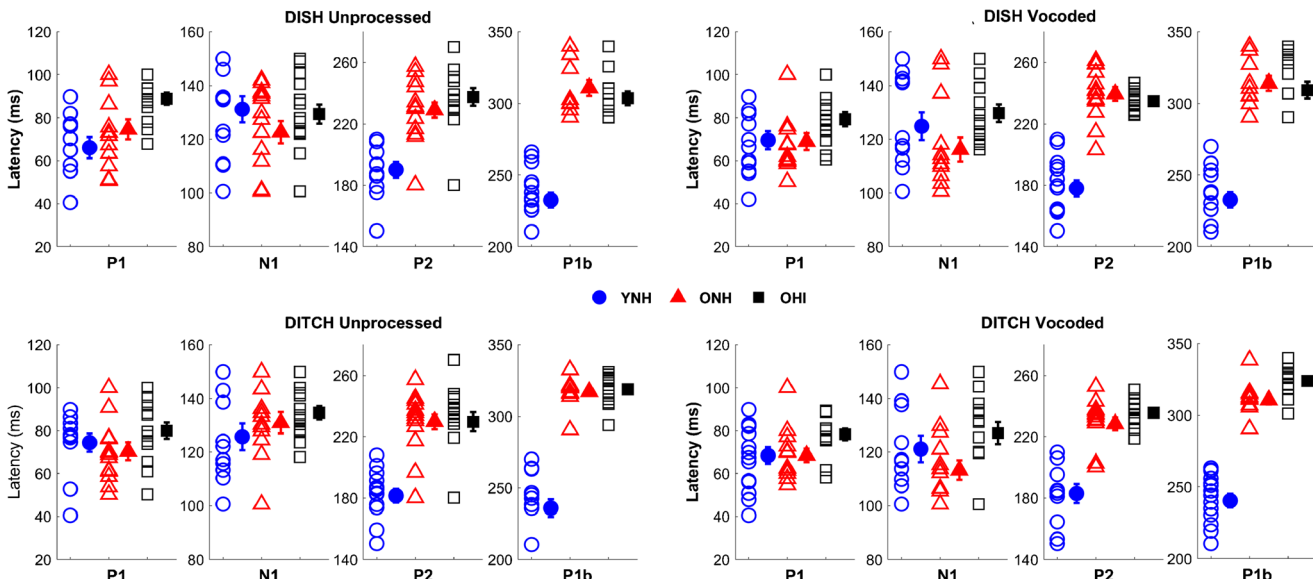
FIG. 10. Split-plot ANOVAs revealed that vocoding decreased N1 latencies across groups ($P < 0.01$), but no vocoding effects were observed for the other peaks. Individual (open symbols) and group mean (closed symbols) for cortical latencies to DISH (top panels) and DITCH (bottom panels) in unprocessed (left column) and vocoded (right column) conditions in young normal-hearing (YNH, blue circles), older normal-hearing (ONH, red triangles), and older hearing-impaired (OHI, black squares) listeners. Error bars show ± 1 standard error

vocoding, and the extent of this change did not differ significantly between groups. Generally, both older groups of listeners had larger amplitudes than the younger listeners for P1, N1, and P1b peaks. No vocoding or group effects were observed for P2 amplitude. Vocoding effects on latency were only observed for N1, such that the latency for vocoded stimuli was earlier than for unprocessed stimuli across groups. In addition, P2 and P1b latencies were earlier for the younger listeners than for either the ONH or OHI listeners.

Controlling for the Effects of Audibility and Age

Controlling for Audibility

As seen in Fig. 1, the ONH listeners have higher thresholds than the YNH listeners at every tested frequency, with greater differences in the high frequencies. Therefore, the aging effects may have been driven, at least in part, by differences in audibility. To reduce the effects of audibility on our electrophysiological results, we covaried for high-frequency pure-tone average (average of thresholds at 4, 6, and 8 kHz) and conducted follow-up analyses of covariance (ANCOVAs) for each result showing a significant difference between the YNH and ONH listeners. The results of the ANCOVAs showed that aging effects persisted for the 50 % crossover point ($F_{(1,27)} = 8.75$, $P = 0.006$, $\eta^2 = 0.24$), DITCH STR correlation ($F_{(1,27)} = 6.15$, $P = 0.02$, $\eta^2 = 0.19$), P2 latency ($F_{(1,25)} = 39.35$, $P < 0.001$, $\eta^2 = 0.61$), and P1b latency ($F_{(1,25)} = 111.37$, $P < 0.001$, $\eta^2 = 0.82$). However, the aging effect did not persist for P1 amplitude ($F_{(1,25)} = 2.87$, $P = 0.10$, $\eta^2 = 0.10$).

As mentioned in the perceptual method, we did not include two of the OHI listeners in the perceptual analysis because they were unable to achieve 90 % accuracy on identification of the perceptual endpoints. We chose to include these two OHI listeners in the EEG analysis, because the stimulus level for the EEG recordings was sufficiently above their thresholds, and we did not want to lose statistical power in our analysis of vocoding effects. When we removed these participants, the effects of age persisted for the cortical effects. However, the FFR (stimulus-to-response correlation) no longer showed a main effect of age when both words (DISH and DITCH) were included, but it did show an effect of age for DITCH only ($P = 0.016$), consistent with our previous study (Roque et al. 2019).

Controlling for Age

We also note that the OHI listeners were significantly older than the ONH listeners ($T_{(1,28)} = 3.08$, $P = 0.005$), so we covaried for age and conducted

ANCOVAs for each result showing significant differences between ONH and OHI listeners. The results of the ANCOVAs showed that hearing loss effects persisted for N1 amplitude ($F_{(1,25)} = 6.76$, $P = 0.016$, $\eta^2 = 0.20$) but not for P1 latency ($F_{(1,25)} = 3.53$, $P = 0.072$, $\eta^2 = 0.13$). Table 2 summarizes results of YNH vs. ONH and ONH vs. OHI comparisons.

Perceptual-Neural Relationships

Multiple Linear Regression. Results of the multiple linear regression analysis indicated that peripheral (PTA) and midbrain (STR) variables significantly contributed to the variance in the crossover point for the vocoded stimuli, but these same variables did not contribute to the slope for the vocoded stimuli (when calculated using PSIGNIFIT and using the linear region of the slope). The "stepwise" method of linear regression was chosen to avoid the bias of order of entry that is present for other methods of linear regression. Independent variables entered into the linear regression included PTA, STR differences between unprocessed and vocoded stimuli for both DISH and DITCH (DISH STR_{DIFF} and DITCH STR_{DIFF}), and P1 amplitude differences between unprocessed and vocoded stimuli for both DISH and DITCH (DISH P1 AMP_{DIFF} and DITCH P1 AMP_{DIFF}). All of these variables were chosen due to observed effects of vocoding and group and to represent different levels of the auditory system (peripheral (PTA), subcortical (STR), cortical (P1 AMP)). The "stepwise" analysis for the slope of the vocoded identification function did not yield any significant models. The "stepwise" analysis for the crossover point yielded two models that were a good fit for the data:

TABLE 2

Summary of analyses of covariance for all significant main effects of group reported in the manuscript

Variable	F value	P value	Effect size
YNH vs. ONH			
CO	8.752	0.006	0.240
STR	6.146	0.02	0.185
P1 AMP	2.870	0.10	0.098
P1b AMP	16.346	< 0.001	0.384
P2 LAT	39.345	< 0.001	0.611
P1b LAT	111.367	< 0.001	0.817
ONH vs. OHI			
N1 AMP	6.762	0.016	0.200
P1 LAT	3.533	0.072	0.128

The YNH vs. ONH comparison covaried for high-frequency pure-tone average and the ONH vs. OHI comparison covaried for age

CO 50 % crossover point, *STR* stimulus-to-response correlation, *P1 AMP* P1 amplitude, *P1b AMP* P1b amplitude, *P2 LAT* P2 latency, *P1b LAT* P1b latency, *N1 AMP* N1 amplitude, *P1 LAT* P1 latency

The first model (PTA ($F_{(1,38)} = 11.29$, $P = 0.002$)) had an $R^2 = 0.23$ and the second model (PTA and DISH P1 AMP_{DIFF} ($F_{(1,38)} = 9.29$, $P < 0.001$)) had an $R^2 = 0.34$. FFR variables did not significantly contribute to the model when including listeners from all groups. Table 3 displays standardized (β) coefficients and levels of significance for the independent variables that were created using the "stepwise" linear regression procedure.

The "stepwise" analysis was then repeated for each group separately to determine if the mechanisms contributing to successful perception of vocoded stimuli differed by group. In this analysis, we reduced the number of variables to three: PTA, DISH STR_{DIFF}, and DISH P1 AMP_{DIFF}. These variables were again chosen to represent three levels of the auditory system (peripheral, midbrain, and cortical). DISH STR_{DIFF} was chosen because it was the only significant variable in the model that included all groups. DISH P1 AMP_{DIFF} was chosen because there was a main effect of group due to larger vocoding-increased changes in amplitude in the ONH and OHI listeners than in the YNH listeners ($F_{(2,39)} = 5.59$, $P = 0.007$, $\eta^2 = 0.22$), but there was no main effect of group for DITCH P1 AMP_{DIFF} ($F_{(2,39)} = 3.07$, $P = 0.06$, $\eta^2 = 0.14$). The analysis of the data from the YNH listeners did not return a model that was a good fit for the data, in that none of the models met the alpha criterion of 0.05 for significance. The analysis of the data from the ONH listeners returned one model that was a good fit for the data (DISH STR_{DIFF} ($F_{(1,12)} = 6.33$, $P = 0.03$)) with an $R^2 = 0.35$. The analysis of the data from the OHI listeners also returned one model that was a good fit for the data (DISH P1 AMP_{DIFF} ($F_{(1,12)} = 11.45$, $P = 0.007$)) with an $R^2 = 0.53$. Table 4 displays standardized (β) coefficients and levels of significance for the independent variables that were created using the "stepwise" analysis in the ONH and OHI groups. Scatter plots showing correlations among these variables are displayed in Fig. 11.

TABLE 3

Standardized (β) coefficients in a model automatically generated by evaluating the significance of each variable's contribution to perceptual crossover of the vocoded identification functions

Variable	R^2 change	β	P value
Model 1	0.23		0.002
PTA		0.48	0.002
Model 2	0.15		0.005
PTA		0.46	0.001
DITCH STR_{DIFF}		-0.39	0.005

Models 1 (PTA) and 2 (PTA and DITCH STRDIFF) contribute significant variance in perceptual crossover point. All other variables were excluded from the model (DISH STRDIFF, DISH, and DITCH P1 AMPDIFF)

TABLE 4

Standardized (β) coefficients in a model automatically generated by evaluating the significance of each variable's contribution to perceptual 50 % crossover of the vocoded identification functions in the older normal-hearing (ONH) and older hearing-impaired (OHI) listeners

Variable	R^2 change	β	P value
ONH listeners			
Model 1	0.45		0.005
DISH STR DIFF		-0.70	0.005
OHI listeners			
Model 1	0.30		0.038
P1 AMP DIFF		0.60	0.038

The variables contributing to the variance in the vocoded crossover point differed by group. In the ONH listeners, the vocoded-related decrease in the DISH stimulus-to-response correlation (DISH STRDIFF) was the only significant factor, and in OHI listeners, the vocoded-related increase in DISH P1 amplitude (P1 AMPDIFF) was the only significant factor. The pure-tone average (PTA) was excluded from the model

DISCUSSION

The purpose of this study was to investigate the effects of spectral degradation on subcortical and cortical responses in YNH, ONH, and OHI listeners and to determine if compensatory mechanisms for this degradation differed between the three groups of listeners. Another purpose was to assess the peripheral and central factors that contribute to perceptual performance for spectrally degraded stimuli. It was hypothesized that spectral degradation would affect perception in ONH and OHI listeners to a greater extent than in YNH listeners. Because of previous evidence that age-related cortical over-representation may compensate for subcortical deficits (Presacco et al. 2019), it was also hypothesized YNH listeners would demonstrate comparatively greater effects of vocoding on FFRs, while the ONH and OHI listeners would demonstrate comparatively greater effects of vocoding on CAEPs.

Consistent with our hypothesis, greater effects of vocoding on the identification functions were seen in ONH and OHI listeners compared to YNH listeners (Fig. 4). For the subcortical measurements, we found strong effects of vocoding on morphology and spectral magnitudes of FFRs across groups. The morphology of the response waveforms (STR correlations) was significantly lower for vocoded than for unprocessed stimuli across groups (Figs. 5 and 6). Analysis of spectral magnitudes in the envelope (F_0) showed significant decreases in responses to vocoded vs. unprocessed stimuli (Fig. 7). Although robust effects of vocoding were found across age groups at the subcortical level, the effects of vocoding were somewhat limited at the cortical level. Vocoding increased P1 amplitude but decreased N1 amplitude in the
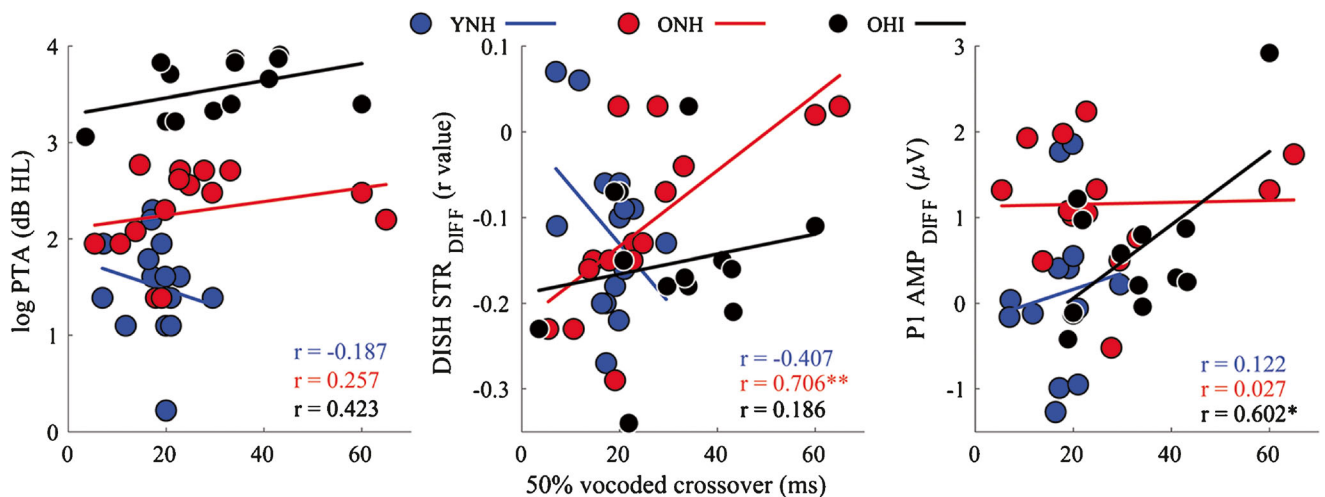
FIG. 11. Reduced effects of vocoding on the FFR to DISH relate to later 50 % crossover points in the ONH listeners, and reduced effects of vocoding on P1 amplitude relate to later crossover points in the OHI listeners. Scatterplots are displayed demonstrating relationships among the vocoded 50 % crossover point and pure-tone average (log PTA), vocoding-induced change in midbrain morphology (DISH STR$_{DIFF}$), vocoding-induced change in P1 amplitude (P1 AMP $_{DIFF}$) in young normal-hearing (YNH, blue), older normal-hearing (ONH, red), and older hearing-impaired (OHI, black) listeners. $*P < 0.05$, $**P < 0.01$

ONH and OHI listeners (Figs. 8 and 9). Vocoding decreased N1 latency across groups (Figs. 8 and 10). Regression models showed that peripheral (PTA) and subcortical (STR correlations) measures were the primary factors that contributed independent variance to perceptual performance (crossover points).

Perceptual

Similar to Goupell et al. (2017), we found that slopes in the linear portion of the psychometric function for vocoded speech were shallower than for unprocessed speech, but only in the OHI listeners. Cortical compensation (Fig. 8) in the YNH listeners may be sufficient to overcome effects of vocoding on the identification function. However, compensation was not complete in the ONH and OHI listeners, as demonstrated by vocoding effects on P1 amplitude and N1 latency in these groups. A combination of reduced compensation and peripheral hearing loss may make the OHI listeners more susceptible to the effects of vocoding on perception than the YNH or OHI listeners.

Across listener groups, we did not find that vocoding significantly affected the crossover point on the DISH-DITCH continuum (Fig. 4). These results contrast with those of Goupell et al. (2017) who found 8-channel vocoding effects in both ONH and YNH listeners. Because we included OHI listeners in this study, we chose to use a stimulus level of 75 dB SPL, which was 10 dB higher than the level used in the Goupell et al. study. Therefore, the boundary of phoneme categorization between the two words was perhaps clearer at the more intense level, even in the vocoded condition.

There was an overall effect of age across stimulus conditions, such that both older groups had later crossover points compared to the YNH group. There were no differences in the crossover points between the older groups. Therefore, temporal processing of silence duration cues appears to be affected primarily by aging rather than significant hearing loss, although it should be noted that some ONH listeners had slight high-frequency hearing loss that may have contributed to the "aging" effect. These results contrast with those of Gordon-Salant et al. (2006), who found that the OHI listeners required longer silence durations to identify DITCH vs. DISH than did ONH listeners. Again, differences between studies might be attributed to differences in presentation levels. Goupell et al. (2017) presented stimuli at 65 dB SPL, we presented stimuli at 75 dB SPL, and Gordon-Salant et al. presented stimuli at 85 dB SPL. Differences across the three studies show a progressive decrease in crossover with an increase in presentation level. Differences in audibility appear to affect the ability to use duration cues, particularly when the stimuli include fricatives and are spectrally degraded.

Subcortical Representation

STR Correlations

There were pronounced effects of vocoding on the morphology of the subcortical response waveforms across groups. Figure 2 shows a reduction in stimulus periodicity with vocoding. The FFR follows the periodicity of the stimulus (Fig. 5), and therefore, a reduction in periodicity with vocoding results in a reduction in the morphology of the FFR. There was

an overall effect of age, with decreased morphology in older compared to younger listeners across unprocessed and vocoded conditions. Although there was a reduction in morphology with vocoding across groups, the effects were more pronounced in the YNH listeners ($\eta^2 = 0.85$) than in the ONH listeners ($\eta^2 = 0.51$) or the OHI listeners ($\eta^2 = 0.76$). These effects are consistent with previous studies that found reduced effects of noise on the FFR in older compared to younger adults (Presacco et al. 2016a; Anderson et al. 2018). Similarly, Mamo et al. (2016) found that distortion of the speech signal through the introduction of jitter resulted in a decline in spectral magnitudes in the YNH but not the ONH listeners. This result is best explained by noting that aging/ hearing loss results in decreased neural synchrony, effectively introducing jitter into the neural signal. Therefore, additional sources of the degradation to the stimulus do not result in further decreases in response magnitude in older adults, whose responses are already closer to the floor.

Spectral Magnitudes

Vocoding significantly decreased the magnitude of the envelope across groups (Fig. 7). These results are consistent with those of Ananthakrishnan et al. (2017), who evaluated FFRs to the vowel /u/ in YNH listeners in different vocoding conditions and found that the decreased periodicity that accompanied fewer vocoder channels resulted in decreases in the F_0 magnitude. However, and more importantly, spectral magnitude was not affected by aging or hearing loss, a similar finding to Roque et al. (2019), who found that the phase-locking factor to a F_0 of 110 Hz did not differ among YNH, ONH, and OHI listeners. Age-related reductions in the envelope may be more pronounced for sustained stimuli (e.g., a prolonged steady-state vowel) due to the loss of auditory nerve fibers that are necessary for sustained neural firing (Presacco et al. 2015).

Cortical Representation

The vocoding effects on CAEPs were much less pervasive than they were on the FFR (Fig. 8). Generally, it appears that the cortex compensates, at least partially, for degraded input from the periphery and midbrain. Previous studies have demonstrated that the cortex can extract information from input that has been significantly altered by neuropathy or other disorders (Kraus et al. 2000; Chambers et al. 2016). However, this compensation ability might not be as robust in older listeners. Generally, the effects of vocoding were more pronounced for the earlier peaks, P1 and N1. The vocoding effect was similar to the

aging effect in that the P1 amplitude was larger in the older than in the younger listeners. These results support previous magnetoencephalography findings that showed that early cortical responses, corresponding to the P1 time region, were over-represented in older compared to younger listeners (Brodbeck et al. 2018). The P1 peak reflects neural encoding of stimulus features (Ceponiene et al. 2005). This peak amplitude may be exaggerated in older listeners because of the necessity to engage a larger neural population to accurately encode a speech stimulus that has been degraded by aging or by vocoding (Reuter-Lorenz and Cappell 2008; Brodbeck et al. 2018; Decruy et al. 2019). Exaggerated amplitudes may also reflect a type of compensatory plasticity that occurs when peripheral input is weakened or distorted (Parthasarathy et al. 2018), but it is unknown whether the time course of compensatory plasticity would be sufficiently rapid to increase firing for vocoded stimuli. We note that the aging effect observed for P1 amplitude did not persist after covarying for high-frequency hearing thresholds; therefore, audibility is likely a main factor in this finding.

In contrast, N1 amplitude decreased for vocoded stimuli compared to unprocessed stimuli. Although the vocoded × group interaction was not significant, the effect sizes were larger in the ONH ($\eta^2 = 0.26$) and OHI ($\eta^2 = 0.15$) listeners compared to the YNH listeners ($\eta^2 = 0.02$), likely because the N1 amplitude in younger listeners is generally smaller than in older listeners. N1 latency was earlier for vocoded stimuli compared to unprocessed stimuli, and again the effect sizes were larger in ONH ($\eta^2 = 0.56$) and OHI ($\eta^2 = 0.45$) listeners than in the YNH listeners ($\eta^2 = 0.06$). The decreased N1 amplitudes to vocoded stimuli may result from expenditure of neural resources for signal detection. Behavioral studies have demonstrated that additional processing needed to understand speech stimuli may expend the resources needed for cognitive functions (Tun et al. 2009; Gosselin and Gagne 2011). The exaggerated amplitudes in older adults' magnetoencephalographic responses occur quite early with a latency of ~30 ms (Brodbeck et al. 2018); therefore, resources may be depleted for the ensuing components of the waveform.

There were no effects of vocoding on the P2 component. Given that P2 is a putative marker of object identification (Ross et al. 2013), it is somewhat surprising that spectral degradation through vocoding did not result in a less robust peak through smaller amplitudes or delayed latencies. However, we note that the endpoints of the vocoded continuum were identified accurately across groups, and therefore, 8-channel vocoding may not have sufficiently degraded the stimulus to result in changes in P2. There was an aging effect on P2 latency; both groups of older

listeners had delayed P2 latencies relative to the younger listeners, consistent with previous cortical aging studies that presented stimuli in quiet or in noise at high signal-to-noise ratios (Tremblay et al. 2003; Billings et al. 2015).

In Fig. 8, a vocoding-related decrease in P1b amplitude is noted for DISH but not for DITCH. The P1b component corresponds to the onset of the fricative. The P2 and P1b components in DISH overlap somewhat; thus, it is possible that vocoding decreased the salience of the final /ʃ/. The 60-ms silence duration prior to the final /ʃ/ in DITCH was apparently sufficient to minimize the effects of vocoding on the P1b component. There were no vocoding effects on the latency of this component. There was an age-related decrease in P1b amplitude and an increase in P1b latency; the latter finding is consistent with the literature showing increased latency delays for the later cortical components (Tremblay et al. 2003; Billings et al. 2015).

Factors Contributing to Behavioral Performance

Linear regression analysis revealed that the degree of vocoding-induced change in subcortical neural responses in midbrain along with peripheral hearing loss predicted the vocoded continuum 50 % crossover point. However, the mechanisms of compensation differ when examining the groups separately. Midbrain representation was the dominant factor in the ONH listeners (change in STR), whereas cortical representation (change in P1 amplitude) was the dominant factor in the OHI listeners. As mentioned above, vocoding may be viewed as disruptive to signal processing in a similar manner to adding external noise to the signal. A greater reduction in morphology with vocoding may suggest that decreased neural resistance to stimulus degradation may lead to poorer perception, at least in the ONH listeners.

In cortex, the extent of compensation (reflected by changes in P1 amplitude) seems to relate to later crossover points with vocoded stimuli. Most of the younger listeners had little if any changes in cortical processing in the vocoded condition and also had the earliest crossover points. A range of change in cortical amplitude was noted in the older listeners, especially those with hearing loss; individuals who had reduced changes to P1 amplitude tended to have earlier crossover points. Older listeners' over-representation of 1-min speech samples in cortex does not correlate with better behavioral performance on sentence recognition in noise (Presacco et al. 2016b). Therefore, poorer performance on the perceptual identification function (later crossover points) may reflect increased cortical compensation for a degraded stimulus that does not benefit actual performance.

These results have clinical implications for cochlear-implant users because they point to a possible neural mechanism for the decreased speech understanding performance for this clinical population (Blamey et al. 2013; Sladen and Zappler 2015). Although the cortex compensates for the degraded (vocoded) speech signal, this compensation mechanism may not be as effective in older cochlear-implant listeners. In future studies, it would be beneficial to evaluate the emergence of categorical perception by recording neural responses to each step of the stimulus continuum, so that the neural response function can be directly compared to the perceptual identification function. This approach was used by Bidelman et al. (2013), who found that categorical perception mapped onto the P2 component of the CAEP in YNH listeners. In addition, it would be interesting to investigate whether neural representation of categorical perception is similar for older acoustic-hearing listeners and for listeners who use cochlear implants.

Limitations

To truly differentiate the effects of aging and hearing loss, it would be optimal to compare age-matched groups of younger and older listeners with and without hearing loss. However, the etiologies and configurations of hearing loss in younger individuals differ greatly from those of older individuals with hearing loss, and therefore, it is difficult to interpret apparent age effects among younger and older listeners with hearing loss. Furthermore, it is quite challenging to recruit older listeners with hearing levels equivalent to those of young listeners, given the normal longitudinal changes in hearing sensitivity with age among women and men, even among a non-noise-exposed sample (Morrell et al. 1996). Therefore, we cannot conclude with certainty that aging affects central encoding of speech stimuli. Nevertheless, we believe that our results provide support for aging effects in that most of the group comparisons demonstrated ONH and OHI listeners had greater effects on neural representation than the YNH listeners, while no differences were found between the ONH and OHI listeners. We would have expected that a hearing loss effect would have resulted in an additional decrement in the OHI compared to the ONH listeners. We also found that most of the aging effects persisted after covarying for high-frequency pure-tone thresholds.

Conclusion

Spectral degradation of a speech signal significantly affects the subcortical neural representation in midbrain for younger and older listeners; however, the

effects on cortical responses are mostly confined to the ONH and OHI listeners. It appears that there is robust compensation in cortex in YNH listeners, which may account for the lack of vocoding effects on the perceptual identification functions in this group. Over-representation of early cortical responses was noted for unprocessed stimuli in the ONH and OHI listeners, and these responses were even more exaggerated for vocoded stimuli. The extent of this exaggeration predicted performance on the perceptual identification function. In the older listeners, it appears that increased neural firing required to encode the vocoded stimulus may reduce available neural resources necessary to accurately perceive the silent duration in the DISH-DITCH contrast.

ACKNOWLEDGMENTS

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest The authors declare that they have no conflicts of interest.

REFERENCES

ANANTHAKRISHNAN S, LUO X, KRISHNAN A (2017) Human frequency following responses to vocoded speech. Ear Hear 38:e256–e267. https://doi.org/10.1097/aud.0000000000000432

ANDERSON S, ELLIS R, MEHTA J, GOUPELL MJ (2018) Age-related differences in binaural masking level differences: behavioral and electrophysiological evidence. J Neurophysiol 120:2939–2952. https://doi.org/10.1152/jn.00255.2018

BENJAMINI Y, HOCHBERG Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol 57:289–300 https://www.jstor.org/stable/2346101

BIDELMAN GM, MORENO S, ALAIN C (2013) Tracing the emergence of categorical speech perception in the human auditory system. NeuroImage 79:201–212. https://doi.org/10.1016/j.neuroimage.2013.04.093

BILLINGS CJ, PENMAN TM, MCMILLAN GP, ELLIS EM (2015) Electro-physiology and perception of speech in noise in older listeners:

effects of hearing impairment and age. Ear Hear 36:710–722. https://doi.org/10.1097/aud.0000000000000191

BLAMEY P, ARTIERES F, BASKENT D, BERGERON F, BEYNON A, BURKE E, DILLIER N, DOWELL R, FRAYSSE B, GALLÉGO S, GOVAERTS PJ, GREEN K, HUBER AM, KLEINE-PUNTE A, MAAT B, MARX M, MAWMAN D, MOSNIER I, O'CONNOR AF, O'LEARY S, ROUSSET A, SCHAUWERS K, SKARZYNSKI H, SKARZYNSKI PH, STERKERS O, TERRANTI A, TRUY E, VAN DE HEYNING P, VENAIL F, VINCENT C, LAZARD DS (2013) Factors affecting auditory performance of postlingustically deaf adults using cochlear implants: an update with 2251 patients. Audiol Neurootol 18:36–47. https://doi.org/10.1159/000343189

BRODBECK C, PRESACCO A, ANDERSON S, SIMON JZ (2018) Over-representation of speech in older adults originates from early response in higher order auditory cortex. Acta Acust United Acust 104:774–777. https://doi.org/10.3813/AAA.919221

CASPARY DM, SCHATTEMAN TA, HUGHES LF (2005) Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: role of inhibitory inputs. J Neurosci 25:10952–10959. https://doi.org/10.1523/jneurosci.2451-05.2005

CEPONIENE R, ALKU P, WESTERFIELD M, TORKI M, TOWNSEND J (2005) ERPs differentiate syllable and nonphonetic sound processing in children and adults. Psychophysiology 42:391–406. https://doi.org/10.1111/j.1469-8986.2005.00305.x

CHAMBERS AR, RESNIK J, YUAN Y, WHITTON JP, EDGE AS, LIBERMAN MC, POLLEY DB (2016) Central gain restores auditory processing following near-complete cochlear denervation. Neuron 89:867–879. https://doi.org/10.1016/j.neuron.2015.12.041

DE CHEVEIGNE A, SIMON JZ (2008) Denoising based on spatial filtering. J Neurosci Methods 171:331–339. https://doi.org/10.1016/j.jneumeth.2008.03.015

DECRUY L, VANTHORNHOUT J, FRANCART T (2019) Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. J Neurophysiol 122:601–615. https://doi.org/10.1152/jn.00687.2018

DELORME A, MAKEIG S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DONG S, MULDERS W, RODGER J, ROBERTSON D (2009) Changes in neuronal activity and gene expression in guinea-pig auditory brainstem after unilateral partial hearing loss. Neuroscience 159:1164–1174. https://doi.org/10.1016/j.neuroscience.2009.01.043

FLORENTINE M, BUUS S, SCHARF B, ZWICKER E (1980) Frequency selectivity in normally-hearing and hearing-impaired observers. J Speech Lang Hear Res 23:646–669. https://doi.org/10.1044/jshr.2303.646

FRIESEN LM, TREMBLAY KL, ROHILA N, WRIGHT RA, SHANNON RV, BASKENT D, RUBINSTEIN JT (2009) Evoked cortical activity and speech recognition as a function of the number of simulated cochlear implant channels. Clin Neurophysiol 120:776–782. https://doi.org/10.1016/j.clinph.2009.01.008

GORDON-SALANT S, FITZGIBBONS PJ (1993) Temporal factors and speech recognition performance in young and elderly listeners. J Speech Hear Res 36:1276–1285. https://doi.org/10.1044/jshr.3606.1276

GORDON-SALANT S, FITZGIBBONS PJ, FRIEDMAN SA (2007) Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners. J Speech Lang Hear Res 50:1181–1193. https://doi.org/10.1044/1092-4388(2007/082)

GORDON-SALANT S, YENI-KOMSHIAN G, FITZGIBBONS P (2008) The role of temporal cues in word identification by younger and older adults: effects of sentence context. J Acoust Soc Am 124:3249–3260. https://doi.org/10.1121/1.2982409

GORDON-SALANT S, FITZGIBBONS PJ, YENI-KOMSHIAN GHAR (2011) Auditory temporal processing and aging: implications for speech

understanding of older people. Audiol Res 1:e4–e4. https://doi.org/10.4081/audiores.2011.e4

GORDON-SALANT S, YENI-KOMSHIAN GH, FITZGIBBONS PJ, BARRETT J (2006) Age-related differences in identification and discrimination of temporal cues in speech segments. J Acoust Soc Am 119:2455–2466. https://doi.org/10.1121/1.2171527

GOSSELIN PA, GAGNE JP (2011) Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. Int J Audiol 50:786–792. https://doi.org/10.3109/14992027.2011.599870

GOUPELL MJ, GASKINS CR, SHADER MJ, WALTER EP, ANDERSON S, GORDON-SALANT S (2017) Age-related differences in the processing of temporal envelope and spectral cues in a speech segment. Ear Hear 38:e335–e342. https://doi.org/10.1097/aud.0000000000000447

HUGHES LF, TURNER JG, PARRISH JL, CASPARY DM (2010) Processing of broadband stimuli across A1 layers in young and aged rats. Hear Res 264:79–85. https://doi.org/10.1016/j.heares.2009.09.005

JASP (2018) JASP Team. In, 0.9 Edition

KOTAK VC, FUJISAWA S, LEE FA, KARTHIKEYAN O, AOKI C, SANES DH (2005) Hearing loss raises excitability in the auditory cortex. J Neurosci 25:3908–3918. https://doi.org/10.1523/jneurosci.5169-04.2005

KRAUS N, BRADLOW MA, CUNNINGHAM CJ, KING CD, KOCH DB, NICOL TG, MCGEE TJ, STEIN LK, WRIGHT BA (2000) Consequences of neural asynchrony: a case of AN. J Assoc Res Otolaryngol 01:33–45. https://doi.org/10.1007/s101620010004

MAMO SK, GROSE JH, BUSS E (2016) Speech-evoked ABR: effects of age and simulated neural temporal jitter. Hear Res 333:201–209. https://doi.org/10.1016/j.heares.2015.09.005

MORRELL CH, GORDON-SALANT S, PEARSON JD, BRANT LJ, FOZARD JL (1996) Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. J Acoust Soc Am 100:1949–1967. https://doi.org/10.1121/1.417906

NASREDDINE ZS, PHILLIPS NA, BÉDIRIAN V, CHARBONNEAU S, WHITEHEAD V, COLLIN I, CUMMINGS JL, CHERTKOW H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53:695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

PARTHASARATHY A, BARTLETT EL (2011) Age-related auditory deficits in temporal processing in F-344 rats. Neuroscience 192:619–630. https://doi.org/10.1016/j.neuroscience.2011.06.042

PARTHASARATHY A, BARTLETT EL, KUJAWA SG (2018) Age-related changes in neural coding of envelope cues: peripheral declines and central compensation. Neuroscience 407:21–31. https://doi.org/10.1016/j.neuroscience.2018.12.007

PHILLIPS SL, GORDON-SALANT S, FITZGIBBONS PJ, YENI-KOMSHIAN G (2000) Frequency and temporal resolution in elderly listeners with good and poor word recognition. J Speech Lang Hear Res 43:217–228. https://doi.org/10.1044/jslhr.4301.217

PICHORA-FULLER MK, SCHNEIDER BA, MACDONALD E, PASS HE, BROWN S (2007) Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. Hear Res 223:114–121. https://doi.org/10.1016/j.heares.2006.10.009

PRESACCO A, SIMON JZ, ANDERSON S (2016A) Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. J Neurophysiol 116:2346–2355. https://doi.org/10.1152/jn.00372.2016

PRESACCO A, SIMON JZ, ANDERSON S (2016B) Effect of informational content of noise on speech representation in the aging midbrain and cortex. J Neurophysiol 116:2356–2367. https://doi.org/10.1152/jn.00373.2016

PRESACCO A, SIMON JZ, ANDERSON S (2019) Speech-in-noise representation in the aging midbrain and cortex: effects of hearing loss. PLoS One 14:e0213899. https://doi.org/10.1371/journal.pone.0213899

PRESACCO A, JENKINS K, LIEBERMAN R, ANDERSON S (2015) Effects of aging on the encoding of dynamic and static components of speech. Ear Hear 36:e352–e363. https://doi.org/10.1097/aud.0000000000000193

REUTER-LORENZ PA, CAPPELL KA (2008) Neurocognitive aging and the compensation hypothesis. Curr Dir Psychol Sci 17:177–182. https://doi.org/10.1111/j.1467-8721.2008.00570.x

ROMERO S, MANANAS MA, BARBANOJ MJ (2006) Quantitative evaluation of automatic ocular removal from simulated EEG signals: regression vs. second order statistics methods. Conf Proc IEEE Eng Med Biol Soc 1:5495–5498. https://doi.org/10.1109/iembs.2006.260338

ROQUE L, GASKINS C, GORDON-SALANT S, GOUPELL MJ, ANDERSON S (2019) Age effects on neural representation and perception of silence duration cues in speech. J Speech Lang Hear Res 62:1099–1116. https://doi.org/10.1044/2018_jslhr-h-ascc7-18-0076

ROSS B, JAMALI S, TREMBLAY KL (2013) Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. BMC Neruosci 14:151. https://doi.org/10.1186/1471-2202-14-151

SÄRELÄ J, VALPOLA H (2005) Denoising source separation. J Mach Learn Res 6:233–272

SCHLÖGL A, KEINRATH C, ZIMMERMANN D, SCHERER R, LEEB R, PFURTSCHELLER G (2007) A fully automated correction method of EOG artifacts in EEG recordings. Clin Neurophysiol 118:98–104. https://doi.org/10.1016/j.clinph.2006.09.003

SCHVARTZ KC, CHATTERJEE M, GORDON-SALANT S (2008) Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. J Acoust Soc Am 124:3972–3988. https://doi.org/10.1121/1.2997434

SHELDON S, PICHORA-FULLER MK, SCHNEIDER BA (2008) Effect of age, presentation method, and learning on identification of noise-vocoded words. J Acoust Soc Am 123:476–488. https://doi.org/10.1121/1.2805676

SLADEN DP, ZAPPLER A (2015) Older and younger adult cochlear implant users: speech recognition in quiet and noise, quality of life, and music perception. Am J Audiol 24:31–39. https://doi.org/10.1044/2014_aja-13-0066

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. https://doi.org/10.1016/S1388-2457(03)00114-7

TUN PA, MCCOY S, WINGFIELD A, AGING (2009) Aging, hearing acuity, and the attentional costs of effortful listening. Psychol Aging 24:761–766. https://doi.org/10.1037/a0014802

WICHMANN FA, HILL NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63:1293–1313. https://doi.org/10.3758/BF03194544

ZHU J, GARCIA E (1999) The Wechsler Abbreviated Scale of Intelligence (WASI). Psychological Corporation, New York