Research Article

Age Effects on Neural Representation and Perception of Silence Duration Cues in Speech

Lindsey Roque,^a Casey Gaskins,^a Sandra Gordon-Salant,^{a,b}
Matthew J. Goupell,^{a,b} and Samira Anderson^{a,b}

Purpose: Degraded temporal processing associated with aging may be a contributing factor to older adults' hearing difficulties, especially in adverse listening environments. This degraded processing may affect the ability to distinguish between words based on temporal duration cues. The current study investigates the effects of aging and hearing loss on cortical and subcortical representation of temporal speech components and on the perception of silent interval duration cues in speech.

Method: Identification functions for the words DISH and DITCH were obtained on a 7-step continuum of silence duration (0–60 ms) prior to the final fricative in participants who are younger with normal hearing (YNH), older with normal hearing (ONH), and older with hearing impairment (OHI). Frequency-following responses and cortical auditory-evoked potentials were recorded to the 2 end points of the continuum. Auditory brainstem responses to clicks were obtained to verify neural integrity and to compare group differences in auditory nerve function. A multiple linear regression analysis was conducted to determine the

peripheral or central factors that contributed to perceptual performance.

Results: ONH and OHI participants required longer silence durations to identify DITCH than did YNH participants. Frequency-following responses showed reduced phase locking and poorer morphology, and cortical auditory-evoked potentials showed prolonged latencies in ONH and OHI participants compared with YNH participants. No group differences were noted for auditory brainstem response Wave I amplitude or Wave V/I ratio. After accounting for the possible effects of hearing loss, linear regression analysis revealed that both midbrain and cortical processing contributed to the variance in the DISH–DITCH perceptual identification functions.

Conclusions: These results suggest that age-related deficits in the ability to encode silence duration cues may be a contributing factor in degraded speech perception. In particular, degraded response morphology relates to performance on perceptual tasks based on silence duration contrasts between words.

Older adults commonly report that hearing difficulty prevents them from fully participating in social situations, leading to social isolation and depression (Heine & Browning, 2002). Although peripheral hearing deficits significantly contribute to speech understanding difficulties, other factors such as central processing and cognitive ability can also affect speech understanding, especially in adverse environments (Humes, Kidd, & Lentz, 2013;

Working Group on Speech Understanding and Aging, 1988). In particular, older adults experience significant declines in auditory temporal processing, and these declines may contribute to decreased speech understanding. Aging affects performance on tasks of lower-level sensory temporal acuity, demonstrated by worse gap detection performance in older adults compared with younger adults (Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994). Higher-level auditory processing of temporal cues is also affected by aging. For example, older adults exhibit worse performance and greater variability on temporal-order tasks of vowel identification in monaural and binaural conditions than do younger adults (Humes, Kewley-Port, Fogerty, & Kinney, 2010). This loss of temporal precision leads to difficulties with speech perception, especially in adverse listening situations (Füllgrabe, Moore, & Stone, 2014; Phillips, Gordon-Salant, Fitzgibbons, & Yeni-Komshian, 2000).

Auditory temporal processing deficits may arise from several age-related physiological changes. Using an aging

^aDepartment of Hearing and Speech Sciences, University of Maryland, College Park

^bNeuroscience and Cognitive Science Program, University of Maryland, College Park

Correspondence to Samira Anderson: sander22@umd.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Steve Aiken

Received March 5, 2018

Revision received June 26, 2018

Accepted August 12, 2018

https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0076

Publisher Note: This article is part of the Special Issue: Select Papers From the 7th Aging and Speech Communication Conference.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

mouse model, Sergeyenko, Lall, Liberman, and Kujawa (2013) observed a reduction in the number of cochlear afferent synapses for low spontaneous rate auditory nerve fibers and a decrease in auditory brainstem response (ABR) Wave I amplitude without a loss of outer hair cells. A decrease in the number of low spontaneous rate fibers was also observed in an aging gerbil model (Schmiedt, Mills, & Boettcher, 1996). These changes would result in a loss of synchronization and possible downstream consequences on the fidelity of signal representation at higher levels of the auditory system from the brainstem to cortex.

In the brainstem, fusiform cells of the cochlear nucleus show reduced phase locking to sinusoidally amplitude-modulated tones in aging rats compared with young rats (Schattman, Hughes, & Caspary, 2008). The authors also noted a change in response types from pauser buildup to wide-chopper responses and suggested that this change is due to loss of inhibitory input. To evaluate aging effects on temporal processing in the midbrain, Walton, Frisina, and O'Neill (1998) used a gap-in-noise paradigm and found a decrease in the number of inferior colliculus (IC) neurons that code very short gap durations and a slower rate of recovery in the neurons of older compared with younger mice. Age-related deficits in neural synchronization have also been observed in auditory cortex. In older rats, there is an overall shift in cortical firing to delayed latencies compared with younger rats (Hughes, Turner, Parrish, & Caspary, 2010). Finally, age-related increases in spontaneous and driven activity in the auditory cortex are observed in a rhesus macaque model (Juarez-Salinas, Engle, Navarro, & Recanzone, 2010) that may be attributed to decreased inhibition associated with reduced effectiveness of GABA neurotransmission in the ascending auditory pathway (Caspary, Ling, Turner, & Hughes, 2008).

These age-related physiological changes may contribute to the temporal processing deficits noted in human studies. Neurophysiological deficits in gap detection observed in the IC of older mice (Walton et al., 1998) may be a factor in poorer gap detection thresholds in older adults compared with younger adults. Schneider and Hamstra (1999) found that older adults demonstrated higher gap detection thresholds compared with younger adults for gap durations of < 500 ms between tonal markers. Older adults have also exhibited increased gap detection thresholds for both speech and non-speech stimuli (Pichora-Fuller, Schneider, Benson, Hamstra, & Storzer, 2006). Larger thresholds in both younger and older adults were noted for stimuli that contained spectrally asymmetrical markers before and after the gap, especially for speech stimuli. These observed age effects for speech stimuli may rely on gap-dependent phonemic contrasts.

For example, the duration of speech segments is an important cue for perception. Older adults, with normal hearing or with hearing loss, have more difficulty than younger adults on tasks that identify words that differ on a temporal dimension, such as silent interval duration to cue the final affricate/fricative (DISH vs. DITCH) or vowel duration to cue final consonant voicing (Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2008; Gordon-Salant, Yeni-Komshian,

Fitzgibbons, & Barrett, 2006; Goupell et al., 2017). Duration-tuned, offset, and other neurons that code temporal properties in the IC (Casseday, Ehrlich, & Covey, 2000; Faure, Fremouw, Casseday, & Covey, 2003) may contribute to the perception of these duration cues.

The frequency-following response (FFR) is an electrophysiological measure that reflects the frequency and temporal characteristics of the stimulus (Moushegian, Rupert, & Stillman, 1973). It arises primarily from the IC with possible contributions from cortex (Coffey, Herholz, Chapesiuk, Baillet, & Zatorre, 2016; Smith, Marsh, & Brown, 1975) and may therefore provide insight into the neural mechanisms underlying older adults' reduced temporal processing of speech. The FFR is particularly well-suited to examine the encoding of a word containing an acoustic temporal cue because of its ability to represent words with such fidelity that recorded waveforms are heard as intelligible speech (Galbraith, Arbagey, Branski, Comerci, & Rector, 1995). Previous FFR studies have demonstrated reduced synchronization to speech and other stimuli in older compared with younger adults (Clinard & Cotter, 2015; Mamo, Grose, & Buss, 2016; Presacco, Jenkins, Liberman, & Anderson, 2015; Presacco, Simon, & Anderson, 2016). A disruption in the synapses between outer hair cells and spiral ganglion cells (Sergeyenko et al., 2013) or a loss of auditory nerve fibers (Schmiedt et al., 1996) would reduce synchronization at the periphery, leading to disruptions in phase locking to the auditory signal.

The cortical auditory-evoked potential (CAEP) can also be used to examine temporal processing in the human listener. CAEPs arise from stimulus-locked postsynaptic potentials of the apical dendrites of pyramidal neurons housed within the auditory cortex (Eggermont, 2001). Stimulus-locked cortical activity may therefore shed light on cortical mechanisms underlying timing and salience of speech processing. The P1–N1–P2 complex is an onset response, indicating detection of an auditory stimulus (Martin, Tremblay, & Stapells, 2007). Specifically, the P1 peak reflects an early pre-perceptual sensory response (Ceponiene, Alku, Westerfield, Torki, & Townsend, 2005; Sharma, Dorman, & Spahr, 2002; Shtyrov et al., 1998), the N1 peak reflects early triggering of attention (Ceponiene, Rinne, & Näätänen, 2002; Näätänen, 1990), and the P2 peak corresponds to early processing of the signal as an auditory object (Näätänen & Winkler, 1999; Ross, Jamali, & Tremblay, 2013). Previous CAEP studies have demonstrated prolonged peak latencies in older adults compared with younger adults, specifically for the P2 latency (Maamor & Billings, 2017; Tremblay, Piskosz, & Souza, 2003). In the Maamor and Billings (2017) study, CAEPs for older adults also exhibited poorer morphology of later-occurring peaks than those for younger adults. The cortical response has also been employed in studies investigating aging effects on neural and behavioral representation of gap detection. Palmer and Musiek (2014) measured thresholds for gaps of 2–20 ms in a 3-s burst of white noise and compared behavioral thresholds to cortical N1–P2 thresholds in younger and older adults with normal hearing. The older adults had higher gap detection thresholds for both

behavioral and neural responses than younger adults, and both groups had close correspondence between behavioral and neural thresholds. Harris, Wilson, Eckert, and Dubno (2012) also assessed gap detection with behavioral testing and CAEPs and observed that cortical peaks were more likely to be present to gaps of a specific duration when behavioral responses were present for that gap duration. They also noted attention effects on gap detection thresholds, in that the older adults who had lower gap detection thresholds also showed decreased P2 latencies and increased N2 amplitudes when they were actively attending to the stimulus.

In the current study, we investigated the effects of age and hearing loss on temporal coding underlying phoneme identification based on silent interval duration. To accomplish this objective, the same stimuli were used for both behavioral and neurophysiological measures, specifically the FFR and CAEP. We hypothesized that aging leads to reduced temporal coding of words containing silent duration intervals, and this loss of temporal coding is reflected in decreased phase locking and morphology in the FFR and prolonged latencies in the CAEP. We also hypothesized that FFR phase locking and CAEP peak latencies contribute to the variance in the 50% crossover point of identification between DISH and DITCH.

Method

Participants

Participants comprised younger adults with normal hearing (YNH, $n = 15$, nine women, six men, 18–26 years, mean age = 21.27 years, $SD = 2.22$), older adults with normal hearing (ONH, $n = 15$, 12 women, three men, 61–78 years, mean age = 68.32 years, $SD = 4.75$), and older adults with hearing impairment (OHI, $n = 15$, eight women, seven men, 63–81 years, mean age = 73.86 years, $SD = 5.10$). There was no significant difference in sex distribution, $\chi^2 = 2.47$, $p = .29$, but the OHI group was significantly older than the ONH group, $t(1,28) = 3.08$, $p = .005$. Clinically normal hearing was defined as hearing thresholds of ≤ 25 dB HL (re: ANSI, 2010) at octave frequencies from 125 to 4000 Hz with no interaural asymmetries of ≥ 15 dB HL at no more than two adjacent frequencies. Participants with hearing loss had pure-tone averages (PTAs) of ≥ 25 dB HL for octave frequencies from 500 to 4000 Hz (see Figure 1). Participants were screened for mild cognitive dysfunction with the Montréal Cognitive Assessment (MoCA), with a cutoff criterion of 22 out of 30 points. We chose to use a lower cutoff than the level typically employed (26), because we included a group of individuals with hearing loss, and hearing impairment may have a negative impact on the MoCA score (Dupuis et al., 2015). Mean MoCA scores and standard deviations of 29.28 ± 0.99 , 26.93 ± 2.52 , and 27.2 ± 1.74 were obtained for YNH, ONH, and OHI participants, respectively. There was a significant main effect of age on the MoCA score, $F(2, 44) = 5.24$, $p = .009$, driven by higher scores in the YNH group than the ONH ($p = .018$)

Figure 1. Mean audiometric thresholds of younger listeners with normal hearing (YNH; blue), older listeners with normal hearing (ONH; red), and older listeners with hearing impairment (OHI; black). Participants with normal hearing have clinically normal hearing (pure-tone thresholds ≤ 25 dB HL at octave frequencies from 125 to 4000 Hz). Participants with hearing impairment have pure-tone averages of ≥ 25 dB HL for octave frequencies from 500 to 4000 Hz. Error bars indicate ± 1 SD. Shaded areas indicate $\pm$ minimum/maximum thresholds.

and OHI groups ($p = .042$), but the two older groups did not differ on the MoCA score ($p = .93$). Other inclusion criteria included an IQ score of ≥ 85 on the Wechsler Abbreviated Scale of Intelligence (Zhu & Garcia 1999). YNH, ONH, and OHI participants obtained mean scores and standard deviations of 115.47 ± 9.55 , 120.45 ± 15.41 , and 111.07 ± 12.99 , respectively, with no significant group differences, $F(2, 44) = 1.61$, $p = .21$. Participants with a history of neurological dysfunction or middle ear surgery were excluded. All participants were native English speakers recruited from the Maryland; Washington, DC; and Virginia areas. All procedures were reviewed and approved by the Institutional Review Board of the University of Maryland, College Park. Participants gave informed consent and were paid for their time.

Stimuli

Test stimuli were the contrasting word pair of speech tokens DISH (483 ms) and DITCH (543 ms) that were first described in Gordon-Salant et al. (2006). Speech tokens were created from isolated recordings of the natural words spoken by an American male adult. This word pair contrast was chosen because it relies on a single acoustic duration cue, the silent interval preceding the final affricate (0 ms) or fricative (60 ms) in DISH and DITCH, respectively. The end point stimulus perceived as DITCH was a hybrid in which the burst and final affricate /tʃ/ were removed from the natural speech token and replaced with the final fricative /ʃ/ from the naturally produced speech token DISH. The final DITCH end point comprised an initial stop,

vowel, closure, and final fricative. The continuum was then generated by removing 10 ms of silence from the closure interval in the DITCH end point until it was 0 ms (the DISH end point). For the perceptual identification functions, participants were presented with the seven steps on the DISH–DITCH continuum of silence duration that ranged from 0 ms (DISH) to 60 ms (DITCH) of silence preceding the final fricative /ʃ/. For the electrophysiology (EEG) recordings, participants were only presented with the two extrema of the DISH–DITCH continuum.

Procedure

Perceptual

The DISH–DITCH continuum was presented to participants in an identification task similar to that implemented in Gordon-Salant et al. (2006). MATLAB (MathWorks, Version 2012a) was used to control the experiment and record participant responses. During the experiment, the participant was seated at a desktop computer in a sound-attenuated booth. The computer monitor displayed three boxes: one that read “Begin Trial” and two side-by-side boxes below that read the words “DISH” and “DITCH.” Testing was self-paced; the participant initiated each trial by clicking the “Begin Trial” box. Stimuli were presented monaurally to the right ear via an insert earphone (ER-2, Etymotic Research) at 75 dB SPL. The participant then identified the stimulus by clicking the DISH or DITCH button on the screen. Participants were encouraged to take a guess if unsure of the stimulus presented and to distribute guesses evenly between DISH and DITCH. Before testing, the participant completed a training run in which only the extrema of the continuum were presented, and he or she was provided feedback. Participants had to achieve 90% accuracy before participating in the experimental runs. The participant then completed five experimental blocks during which he or she did not receive feedback. Stimuli along the DISH–DITCH continuum were each presented a total of 10 times over the course of the experimental runs. It should be noted that two OHI participants did not complete the perceptual task because they could not meet the criteria of the training procedure after five attempts.

EEG

Prior to the EEG recording, the final fricative /ʃ/ was excised from DISH, and thresholds were obtained to the /ʃ/ presented monaurally to the right ear. This was done to ensure that the high-frequency regions of the stimuli used in FFR/CAEP testing were audible. The stimuli were delivered via Presentation software (Neurobehavioral Systems, Inc.) to a single insert earphone (ER-1, Etymotic Research) while the participant was seated in the EEG recording booth. Two participants were excluded because they did not meet the threshold criterion of < 55 dB SPL for the /ʃ/ stimulus. The first formant of the /ʃ/ region was approximately 2000 Hz and had an intensity level of 66 dB SPL. This level would correspond to 59.5 dB HL and is above

the highest threshold at 2000 Hz (55 dB HL) in the OHI group.

FFR. The end point stimuli from the DISH and DITCH continuum were presented monaurally to the right ear via Presentation software through an insert earphone (ER-1, Etymotic Research) at 75 dB SPL. The levels of the individual components were 76 and 66 dB SPL for the vowel /ɪ/ and fricative /ʃ/, respectively. The entire DISH and DITCH waveforms were presented with alternating polarities at a rate of 1.5 Hz. The presence of stimulus artifact was evaluated by presenting the stimulus through an occluded insert earphone. In addition, each response waveform was examined to ensure that the response onset was consistent with the expected delays of the auditory system (~6–8 ms). FFRs were recorded at a sampling rate of 16,384 Hz using the BioSemi ActiABR200 acquisition system (BioSemi B.V.). A standard vertical montage of five electrodes (Cz, two forehead ground CMS/DRL electrodes, two earlobe reference electrodes) was used. A minimum of 3,000 artifact-free sweeps were recorded for each speech token from each participant.

Cortical. CAEPs were also recorded to DISH and DITCH stimuli presented at a rate of 0.71 Hz at 75 dB SPL. Responses were recorded at a sampling rate of 2,048 Hz using the BioSemi ActiveTwo system via a 32-channel electrode cap with average earlobes (A1 and A2) serving as references. A minimum of 500 artifact-free sweeps were recorded for each speech token from each participant.

Click ABR. Click-evoked ABRs to 100-μs click stimuli were obtained from each ear to verify neural integrity and to obtain an indirect measure of auditory nerve function. Click stimuli were presented monaurally via insert earphones (ER-3A, Etymotic Research) at 80 dB SPL (Intelligent Hearing System SmartEP system, IHS). A two-channel vertical montage of four electrodes (Cz, one forehead ground electrode, two earlobe reference electrodes) was used. Two sets of 2,000 sweeps were obtained at a rate of 21.1 Hz for each ear.

Recordings took place over two sessions: FFRs were recorded during one session, and click-evoked ABRs and CAEPs were recorded during the other. During each recording session (~1.5 hr), participants sat in a reclining chair and watched a silent, captioned movie of their choice to facilitate a relaxed yet wakeful state.

Data Analysis

Perceptual

Percent identification of DISH responses was determined for each step on the continuum. Slope and 50% crossover point (boundary of stimulus categorization) were calculated from the identification functions using the same procedure performed in Gordon-Salant et al. (2006). Linear regression was performed on data points on the linear portion of the identification function, which fell between approximately 20% and 80% identification of DISH. The 50% crossover point was then calculated from the following formula where values *a* and *b* represent slope and *y*-intercept, respectively: $x = (50 - b)/a$.

EEG

FFR data reduction. Data recorded with BioSemi were analyzed in MATLAB (MathWorks, Version R2011b) after being converted into MATLAB format using the function `pop_biosig` from EEGLab (Delorme & Makeig, 2004). Sweeps with amplitude in the $\pm 30 \mu\text{V}$ range were retained, averaged in real time, and then processed offline. Recordings underwent an offline bandpass filter of 70–2000 Hz using a zero-phase, fourth-order Butterworth filter. To maximize the response to the temporal envelope, a final average response was created by averaging the sweeps of both polarities over a time window of 660 ms. To maximize the response to the temporal fine structure, a final average response was created by subtracting sweeps of one polarity from the other.

Stimulus-to-response correlations. Stimulus-to-response (STR) correlations examined the degree to which the envelope of the response waveforms mimicked those of the stimulus waveforms for DISH and DITCH. The envelope of the stimulus was extracted and bandpass filtered with the same filter that was used for the response. The STR correlation r value was obtained using the cross-correlation function in MATLAB by shifting the stimulus waveform in time relative to the response waveform until a maximum correlation was found from 10 to 300 ms.

Phase-locking factor. Phase-locking factor (PLF) evaluates trial-by-trial phase coherence for a specific frequency range at each individual point in time during a response. Accurate phase tracking to the speech envelope may be an important contributor to speech intelligibility (Giraud & Poeppel, 2012; Luo & Poeppel, 2007), and previous studies have shown that older adults have reduced phase locking (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2012; Ruggles, Bharadwaj, & Shinn-Cunningham, 2012), especially to a sustained vowel (Presacco et al., 2015). Complex Morlet wavelets were used to investigate phase locking to optimize the frequency resolution at all tested frequencies using procedures described in previous studies (Jenkins, Fodor, Presacco, & Anderson, 2017; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). The wavelet was convolved with the response at each discrete time point, yielding an amplitude (arbitrary units) and phase value (radians) for each Time $\times$ Frequency bin. The amplitude was normalized to a value of 1 irrespective of voltage, and the resulting normalized energy was calculated for each sweep and then averaged across trials. The wavelets were used to decompose the signal from 80 to 800 Hz, which were used to calculate PLF to the temporal envelope (PLF_{ENV}). Individual PLF_{ENV} values to the fundamental frequency (F_0) of the stimulus vowel /i/ (110 Hz) were calculated and averaged for each group. Wavelets were also used to decompose the signal from 300 to 1500 Hz, which were used to calculate the PLF to the temporal fine structure (PLF_{TFS}). Individual PLF_{TFS} values to the first formant of /i/ (400 Hz) were calculated and averaged for each group. For both stimuli, PLF_{ENV} and PLF_{TFS} were calculated for the vowel region (50–190 ms) to examine different features of the vowel (F_0 and F_1), as it was assumed that increased neural

representation of the vowel region would enhance perception of the presence of the silent duration.

Cortical. Recordings underwent an offline bandpass filter of 1–30 Hz using a zero-phase, fourth-order Butterworth filter. A regression-based electrooculography reduction method was used to remove eye movements from the filtered data (Romero, Mañanas, & Barbanj, 2006; Schlögl et al., 2007). For each sweep, the time window –500 to 1000 ms was referenced to the stimulus onset. A final response was averaged from the first 500 artifact-free sweeps. The denoising source separation algorithm was used to obtain artifact-free data from the 32 recorded channels (de Cheveigne & Simon, 2008; Särelä & Valpola, 2005). An automated peak-picking algorithm was implemented in MATLAB to identify the latencies for prominent cortical peaks in their expected time regions: P1 (40–90 ms), N1 (100–150 ms), P2 (150–210 ms), and P1b (210–270 ms) in YNH participants and P1 (50–100 ms), N1 (100–150 ms), P2 (180–270 ms), and P1b (290–340 ms) in ONH and OHI participants. These expected latency ranges were determined from the average response for the YNH, ONH, and OHI groups. The same MATLAB algorithm calculated the area amplitudes corresponding to these time ranges. Note that the P1b peak corresponds to the onset of the /f/.

Click ABR. Click-evoked ABRs were offline bandpass filtered from 70 to 2000 Hz using a zero-phase, sixth-order Butterworth filter. The total 4,000 sweeps for each ear were averaged. An automated peak-picking algorithm in MATLAB was used to identify peak latencies and amplitudes for Waves I, III, and V that were closest to the expected latency (within 0.5 ms), based on average latencies obtained in Anderson et al. (2012). Each peak identification was confirmed by a trained peak picker. Changes were made where appropriate. Two measures were calculated from the averaged click responses: Wave I amplitude and Wave V/I amplitude ratio. A vertical montage was used to determine the Wave V/I amplitude ratio, as has been used in previous investigations of potential synaptopathy in humans (Grose, Buss, & Hall, 2017; Guest, Munro, Prendergast, Millman, & Plack, 2018). A derived horizontal montage was used to maximize Wave I amplitude.

Statistical Analysis

All statistical analyses were conducted in SPSS Version 23.0. Repeated-measures analyses of variance were used to examine effects of group (YNH vs. ONH vs. OHI) and within-group effects of stimulus (DISH vs. DITCH) on FFR variables (STR, PLF_{ENV} , and PLF_{TFS}) in the time region preceding the silent interval and CAEP variables (peak latencies and amplitudes). Univariate one-way analyses of variance were used for group comparisons of perceptual 50% crossover points and slope, ABR Wave V/I ratio, and ABR Wave I amplitude. Scheffé's post hoc analyses were used when significant main effects or interactions were found. Fisher's z -transform was applied to the STR r value before statistical analysis. The Levene Test was used to ensure homogeneity of all measures. Pearson's

correlations were used to assess relationships among perceptual, peripheral, cortical, and midbrain measures. The false discovery rate was used to correct for multiple comparisons (Benjamini & Hochberg, 1995). Linear regressions were performed with 50% crossover point entered as the dependent variable and PTA (average of right ear threshold for 500, 1000, 2000, 3000, and 4000 Hz), STR r value, PLF_{TFS}, CAEP N1 amplitude, and P2 latency to DITCH entered as independent variables in that order. The order of variable entry using the “Enter” method of multiple linear regression was specified to mirror the ascension of incoming acoustic signals through the auditory system. In all models, the residuals for normality were checked to ensure that the linear regression analysis was appropriate for the data set. Collinearity diagnostics were run with satisfactory variance inflation factor (highest = 1.64) and tolerance (lowest = 0.61) scores, indicating the absence of strong correlations between two or more predictors.

Results

Perceptual

The average identification functions for YNH, ONH, and OHI participants are shown in Figure 2. To quantify participants' boundary of phonemic categorization, 50% performance reduction points and slope were calculated for perceptual judgments along the DISH–DITCH continuum of silence duration. Mean crossover points and standard errors for YNH, ONH, and OHI participants were $14.43 \pm$

Figure 2. Average identification functions for percentage of trials identified as DISH as a function of silence duration for the three participant groups. Older participants with normal hearing (ONH) and hearing loss (OHI) required longer silence durations to discriminate DISH–DITCH than did younger participants (YNH). Error bars indicate ± 1 SE.

0.83, 22.30 ± 0.99 , and 20.59 ± 1.29 , respectively. Mean slopes and standard errors for YNH, ONH, and OHI participants were -0.55 ± 0.06 , -0.49 ± 0.05 , and -0.41 ± 0.07 , respectively. A main effect of group was observed for the 50% crossover point, $F(2, 41) = 16.68$, $p < 0.001$, $\eta_p^2 = .46$. Post hoc analyses showed that younger listeners exhibited earlier phonemic categorization boundaries relative to both groups of older listeners (all $ps < .05$). No group differences were observed between the two groups of older listeners ($p = .52$). There was no effect of group on slope of the identification functions, $F(2, 41) = 1.91$, $p = .16$, $\eta_p^2 = .09$.

EEG

FFR

STR. Figure 3 compares average response waveforms of the three participant groups to the stimuli spectra and waveforms. There were no significant effects of group, $F(2, 42) = 2.96$, $p = .06$, $\eta_p^2 = .12$ or stimulus, $F(1, 42) = 0.63$, $p = .43$, $\eta_p^2 = .02$. The Stimulus $\times$ Group interaction was not significant, $F(2, 42) = 1.88$, $p = .17$, $\eta_p^2 = .08$. Post hoc analyses were performed to further examine group differences because of the observed p value (.06) for the effect of group that was close to the criterion α level of .05. No differences between the three groups were observed for DISH, $F(2, 44) = 0.021$, $p = .447$, $\eta_p^2 = .038$, but a main effect of group was observed for DITCH, $F(2, 44) = 6.028$, $p = .005$, $\eta_p^2 = .223$. Response waveforms of YNH listeners were more closely correlated with the stimulus waveform than the response waveforms of either the ONH ($p = .008$) or OHI listeners ($p = .047$). Differences between the two older listener groups were not significant ($p = .762$).

Envelope PLF. Figure 4 compares average phase locking to the temporal envelopes of the DISH and DITCH stimuli across the three participant groups. For both stimuli, PLF_{ENV} values to the F0 (110 Hz) were calculated for the vowel region (50–190 ms). There were no effects of group, $F(2, 42) = 0.44$, $p = .65$, $\eta_p^2 = .02$ or stimulus, $F(1, 42) = 0.04$, $p = .85$, $\eta_p^2 < .01$, and the Stimulus $\times$ Group interaction was not significant, $F(2, 42) = 0.15$, $p = .86$, $\eta_p^2 < .01$.

Fine structure PLF. Average phase locking to the temporal fine structure of DISH and DITCH stimuli for the three participant groups is displayed in Figure 5. PLF_{TFS} values to 400 Hz (first formant) were calculated in the 140-ms vowel region. There was a main effect of group, $F(2, 42) = 5.29$, $p = .01$, $\eta_p^2 = .20$. Post hoc analyses revealed that YNH participants exhibited more robust phase locking compared with ONH participants ($p = .01$), but not OHI participants ($p = .09$). No differences were observed between the two groups of older participants ($p = .69$). No effect of stimulus was observed, $F(1, 42) = 3.72$, $p = .06$, $\eta_p^2 = .08$, but there was a significant Stimulus $\times$ Group interaction, $F(2, 42) = 4.23$, $p = .02$, $\eta_p^2 = .17$. This interaction was driven by a main effect of group for DITCH, $F(2, 42) = 6.90$, $p = .003$, $\eta_p^2 = .25$, but not for DISH, $F(2, 42) = 2.25$, $p = .12$, $\eta_p^2 = .10$. As observed in Figure 5, YNH adults had higher PLF_{TFS} values for DITCH compared with DISH

Figure 3. Spectra (A) and waveforms (B) for DISH (0-ms silence duration; left column) and DITCH (60-ms silence duration; right column) stimuli. (C) Average response waveforms in the time domain to DISH and DITCH for younger listeners with normal hearing (YNH; blue), older listeners with normal hearing (ONH; red), and older listeners with hearing impairment (OHI; black). The periodicity of the stimuli, representing the 9.1-ms period of the 110-Hz F0, is mirrored in the average response waveforms. No group differences were observed to the speech token DISH. Younger listeners, however, demonstrated better approximations to the stimulus waveform for DITCH than did either group of older listeners. (D) Individual (open symbol) and group average (closed symbol) stimulus-to-response correlation r values displayed for DISH and DITCH. Error bars indicate ± 1 SE.

($p = .02$), but no stimulus differences were observed in the ONH or OHI groups (all $ps > .05$).

Cortical

Amplitude. Figure 6 displays individual and average CAEP response waveforms obtained from the denoising source separation analysis and Cz electrode for the three participant groups. Amplitudes were quantified for major components of measured CAEPs. Main effects of peak, $F(3, 39) = 14.98$, $p < .001$, $\eta_p^2 = .54$, and group, $F(2, 41) = 5.56$, $p = .01$, $\eta_p^2 = .21$, were observed on CAEP peak amplitude. Significant Peak $\times$ Group, $F(6, 80) = 11.79$, $p < .001$, $\eta_p^2 = .47$, Stimulus $\times$ Peak, $F(3, 39) = 7.05$, $p = .001$, $\eta_p^2 = .35$, and Stimulus $\times$ Peak $\times$ Group interactions, $F(6, 80) = 4.24$, $p < .001$, $\eta_p^2 = .24$, were also observed. OHI adults exhibited greater N1 peak amplitudes compared with both YNH ($p = .01$) and ONH adults ($p = .05$). CAEP peak amplitudes also exhibited greater variability for the DITCH stimulus than DISH. The Stimulus $\times$ Peak $\times$ Group interaction was driven by the increased P1b peak amplitude exhibited by younger adults compared with older adults that was observed for DISH but not for DITCH speech tokens. There was neither a significant main effect of stimulus, $F(1, 41) = 2.85$,

$p = .10$, $\eta_p^2 = .07$, nor Stimulus $\times$ Group interaction, $F(2, 41) = 1.33$, $p = .27$, $\eta_p^2 = .06$.

Latency. Figure 7 displays individual and average amplitudes and latencies for prominent CAEP peaks for the three participant groups obtained from the denoising source separation analysis. A main effect of peak was observed, $F(3, 39) = 2,163.68$, $p < .001$, $\eta_p^2 = .99$, such that latency was greater with later-occurring CAEP P2 and P1b peak components. There was also a significant main effect of group, $F(1, 41) = 0.82$, $p = .37$, $\eta_p^2 = .02$. Significant Peak $\times$ Group, $F(6, 80) = 10.20$, $p < .001$, $\eta_p^2 = 0.43$, Stimulus $\times$ Peak, $F(3, 39) = 3.49$, $p = .03$, $\eta_p^2 = .21$, and Stimulus $\times$ Peak $\times$ Group interactions, $F(6, 80) = 2.43$, $p = .03$, $\eta_p^2 = .15$, were also observed. Post hoc analyses revealed that younger participants exhibited earlier peak latencies than both groups of older participants (all $ps < .001$), whereas no difference was observed between latencies for ONH and OHI participants ($p = .16$). Such group differences are particularly pronounced for the later-occurring P2 and P1b CAEP components. The Stimulus $\times$ Peak $\times$ Group interaction was driven by the prolonged P1 latency exhibited by OHI participants compared with that of the two groups of participants with normal hearing that was observed for

Figure 4. A: Average phase-locking factor (PLF) to the temporal envelope (ENV) of DISH and DITCH stimuli represented in the time-frequency domain, with hotter (red) colors indicating increased PLF. B: Individual (open symbol) and group average (closed symbol) PLF to 110 Hz (F0) of the stimulus in time region corresponding to the vowel /i/ (50–190 ms). No group differences were noted for either word. Error bars indicate ± 1 SE. YNH = younger listeners with normal hearing; ONH = older listeners with normal hearing; OHI = older listeners with hearing impairment.

Figure 5. A: Average phase-locking factor (PLF) to the temporal fine structure (TFS) of DISH and DITCH stimuli represented in the time-frequency domain, with hotter colors (red) indicating increased PLF. B: Individual (open symbol) and group average (closed symbol) PLF to 400 Hz (F1) of the stimulus in time region corresponding to the vowel /i/ (50–190 ms). Younger listeners (YNH) exhibited more robust phase locking compared with older listeners with normal hearing (ONH) and hearing loss (OHI). Error bars indicate ± 1 SE.

Figure 6. Average cortical auditory-evoked potentials obtained through the denoising source separation algorithm (DSS; top) and for single Cz electrode (bottom) for younger listeners with normal hearing (YNH; blue), older listeners with normal hearing (ONH; red), and older listeners with hearing impairment (OHI; black).

Figure 7. Top: Individual (open symbol) and group average (closed symbol) amplitude for prominent cortical peaks to DISH (left) and DITCH (right) obtained from the denoising source separation algorithm for the three participant groups. Older adults with hearing loss (OHI) exhibited increased N1 amplitudes compared with both normal hearing groups to both DISH and DITCH. Bottom: Individual (open symbol) and group average (closed symbol) latencies for prominent cortical peaks to DISH (left) and DITCH (right) obtained from the denoising source separation algorithm for the three participant groups. Older listeners with hearing loss exhibited longer P1 latencies than did younger (YNH) and older listeners with normal hearing (ONH) to DISH. N1 latencies are equivalent across groups to both speech tokens. Both groups of older listeners exhibited larger peak latencies for P2 and P1b than did younger listeners to both speech tokens. Error bars show ± 1 SE.

DISH but not for DITCH. There was no significant effect of stimulus, $F(1, 41) = 0.82, p = .37, \eta_p^2 = .02$, and no Stimulus $\times$ Group interaction, $F(2, 41) = 0.73, p = .49, \eta_p^2 = .04$.

Click ABR

Figure 8 displays average click-evoked ABRs for the three participant groups. Wave I amplitudes of recorded click-evoked ABRs were calculated to verify the integrity of the vestibulocochlear nerve in each ear and to obtain an indirect measure of cochlear synaptopathy. No effect of group was observed on Wave I amplitude, $F(2, 41) = 2.87, p = .07, \eta_p^2 = .13$. Amplitude ratios between Waves V and I were also calculated. No differences in Wave V/I ratio were observed between groups, $F(2, 41) = 0.82, p = .45, \eta_p^2 = .04$.

Perceptual–Neural Relationships

Multiple regression. Results of multiple linear regression analysis indicated that both midbrain and cortical variables predicted variance in the 50% perceptual crossover points beyond the contribution of PTA (see Table 1). The “Enter” method of linear regression was chosen because it specifies the order in which variables are entered into the model. The regression was repeated hierarchically with PTA on the first step, DITCH STR correlation and PLF_{TFS}

on the second step, and DITCH N1 amplitude and P2 latency on the third step. These variables were selected based on demonstrated group differences and to represent different levels of the auditory system. P2 latency was chosen instead of P1b latency because of an observed greater spread of data. We did not include the ABR variables in the regression because we wanted to focus on measures using the same stimuli.

It was hypothesized that FFR and CAEP variables would predict the greatest amount of variance in perceptual performance. By entering PTA first, the possibility that FFR and CAEP variables would account for the variance that might have otherwise been attributed to peripheral auditory deficits was reduced. The final model (PTA, STR, PLF_{TFS}, N1 amplitude, and P2 latency entered in this order) is a good fit for the data, $F(5, 36) = 4.65, p = .002$, with an R^2 value of .39. In Model 3, only P2 latency significantly contributed to the variance in the 50% crossover point. A “Stepwise” method of linear regression was then performed to avoid the bias of order of entry used in the hierarchical linear regression. This model is also a good fit for the data, $F(2, 39) = 11.02, p < .001$, with an R^2 value of .36. PTA, DITCH PLF_{TFS}, and N1 amplitude were excluded from the final model because they did not significantly contribute

Figure 8. Top: Average auditory response waveforms to a 100- μ s click in younger listeners with normal hearing (YNH; blue), older listeners with normal hearing (ONH; red), and older listeners with hearing impairment (OHI; black) participants using a vertical montage (left) and a derived horizontal montage (right). The major peaks are labeled as I, III, or V. Bottom: The Wave V/I ratio from the vertical montage (left) and the Wave I amplitude from the derived horizontal montage (right) are quantified. No significant group differences were found, but considerable individual variability was observed as can be seen in the large error bars. Error bars indicate ± 1 SE.

Table 1. Summary of “Enter” hierarchical regression analysis for variables.

Variable	R^2 change	B	SE	β	95% CI	p
Model 1	.08					.076
PTA		0.02	0.01	0.28	−0.00, 0.04	.023
Model 2	.16					.028
PTA		0.01	0.01	0.01	−0.01, 0.03	.291
DITCH STR		−2.28	1.85	−2.13	−6.01, 1.47	.226
DITCH PLF _{TFS}		−40.98	28.50	−0.26	−98.67, 16.70	.159
Model 3	.16					.016
PTA		0.00	0.01	0.01	−0.02, 0.03	.767
DITCH STR		−2.16	1.69	−2.02	−5.59, 1.28	.211
DITCH PLF _{TFS}		−31.282	26.54	−0.20	−85.10, 22.54	.246
DITCH N1 AMP		−0.11	0.01	−0.13	−0.40, 0.17	.431
DITCH P2 LAT		0.02	0.14	0.45	0.01, 0.03	.005

Note. Unstandardized (B) standard error (SE) and standardized (β) coefficients in three models comparing separate contributions from peripheral (PTA), midbrain (DITCH STR and DITCH PLF_{TFS}), and cortical variables (DITCH N1 AMP and DITCH P2 LAT). Model 1 (peripheral) does not significantly predict variance in the 50% crossover point. Models 2 (midbrain) and 3 (cortical) contribute significant additional variance to the previous models in the crossover point. DITCH P2 LAT accounts for the greatest amount of variance in the model. CI = confidence interval; PTA = pure-tone average; STR = stimulus-to-response correlation; PLF_{TFS} = phase-locking factor to the temporal fine structure; N1 AMP = N1 amplitude; P2 LAT = P2 latency.

to the variance. Table 1 displays unstandardized B and standard deviations, standardized (β) coefficients, confidence intervals, and levels of significance for the independent variables for the three models that were created using the hierarchical regression procedure. Table 2 displays unstandardized B and standard deviations, standardized (β) coefficients, confidence intervals, and levels of significance for the independent variables for the two models that were created using the stepwise regression procedure. Table 3 displays correlational values among the independent variables used in the linear regression. Figure 9 displays scatter plots of relationships among perceptual and neural variables.

Discussion

The main purpose of this study was to achieve a better understanding of potential factors underlying speech perception deficits in older adults, particularly the ability to use temporal cues for phoneme identification. To that end, we investigated the effects of age and hearing loss on temporal coding underlying phoneme identification based

on silent interval duration, and we assessed the possible neural contributions to variability in perceptual performance on the phoneme identification task. Older participants required longer silence durations to perceive the final affricate in DITCH versus the fricative in DISH than did younger participants, consistent with previous behavioral experiments (Gordon-Salant et al., 2006; Goupell et al., 2017). Electrophysiological measurements revealed age-related deficits in neural representation of the end points of the DISH–DITCH continuum. In the FFR, poorer morphology (lower STR correlations) was observed in the response waveforms of older participants compared with younger participants, but only for the word DITCH, which contained the 60-ms silence interval. Both older groups also showed reduced phase locking to the temporal fine structure compared with younger participants, and these effects were again specific to the word DITCH. In cortical responses, older participants exhibited increased amplitudes for earlier-occurring peaks and prolonged latencies for later-occurring peaks compared with younger participants. No group differences were found for Wave I amplitude or Wave V/I ratio in the

Table 2. Summary of “Stepwise” regression analysis for variables.

Variable	R^2 change	B	SE	β	95% CI	p
Model 1	.28					< .001
DITCH P2 LAT		0.02	0.01	0.53	0.01, 0.03	< .001
Model 2	.08					.032
DITCH P2 LAT		0.02	< 0.01	0.47	0.01, 0.03	.001
DITCH STR		−3.12	1.40	−0.29	−5.95, −0.29	.032

Note. Unstandardized (B) and standard error (SE) coefficients and standardized (β) coefficients in two models automatically generated by evaluating the significance of each variable’s contribution to the 50% crossover point. Three models were generated. Model 1 (DITCH P2 LAT) predicts significant variance in the 50% crossover point. The addition of DITCH STR in Model 2 (midbrain) predicts significant additional variance, but all other variables were excluded from the model (PTA, DITCH PLF_{TFS}, and DITCH N1 AMP; $N = 43$). CI = confidence interval; P2 LAT = P2 latency; STR = stimulus-to-response correlation; PTA = pure-tone average; PLF_{TFS} = phase-locking factor to the temporal fine structure; N1 AMP = N1 amplitude.

Table 3. Intercorrelations among the 50% crossover point and the independent peripheral, midbrain, and cortical variables.

Variables	1. Crossover	2. PTA	3. DITCH STR	4. DITCH PLF _{TFS}	5. DITCH N1 AMP	6. DITCH P1 LAT
1. Crossover						
2. PTA	.27					
3. DITCH STR	-.38*	-.27				
4. DITCH PLF _{TFS}	-.42*	-.30	.58***			
5. DITCH N1 AMP	.16	.55***	-.25	-.32*		
6. DITCH P2 LAT	-.53**	.36*	-.15	-.27	.30	

Note. Results of Pearson’s correlational analysis indicated that the 50% crossover is most strongly correlated to P1 LAT. However, midbrain variables (STR and PLF_{TFS}) are also correlated with the 50% crossover. The levels of significance reflect corrections for multiple comparisons. PTA = pure-tone average; STR = stimulus-to-response correlation; PLF_{TFS} = phase-locking factor to the temporal fine structure; N1 AMP = N1 amplitude; P1 LAT = P1 latency; P2 LAT = P2 latency.

* $p < .05$. ** $p < .01$. *** $p < .001$.

click ABRs. Linear regression analysis revealed that both midbrain and cortical variables made significant contributions to variance in the 50% perceptual crossover point.

Perceptual

The lack of difference between perceptual crossover points of older participants with normal hearing and with hearing loss suggests that age, rather than hearing loss, may be driving group differences in performance (see Figure 2). These age effects are consistent with those observed in Gordon-Salant et al. (2006). However, Gordon-Salant

et al. (2006) found that older participants with hearing loss had later 50% perceptual crossover points than older participants with normal hearing. We did not observe similar group differences in the current study. Because we found a significant age difference between the two older groups, we repeated the analysis and restricted the age range to > 66 years, increasing the mean age of the nine remaining ONH participants to 71.27 (± 3.66 SD) and of the remaining 14 OHI participants to 74.58 (± 4.44 SD) and reducing age difference, $t(1,21) = 1.86$, $p = .08$. There was no change in the overall pattern from the results with the larger number of participants. There was a significant main effect

Figure 9. Scatter plots demonstrating relationships among perceptual (50% crossover point on the perceptual identification function), peripheral (PTA), and neural variables (DITCH STR, DITCH PLF TFS, DITCH N1 amplitude, and DITCH P2 latency). Significant correlations were obtained between the 50% crossover point and DITCH STR, DITCH PLF TFS, and DITCH P2 latency. PTA = pure-tone average; STR = stimulus-to-response correlation; PLF TFS = phase-locking factor to the temporal fine structure.

of group, $F(2,36) = 17.39$, $p < .001$, $\eta_p^2 < .51$, with earlier crossover points in YNH participants compared with either ONH participants ($p < .001$) or OHI participants ($p = .001$), and no differences between the older groups ($p = .21$).

The differences in findings between the current study and those of Gordon-Salant et al. (2006) may be attributed to differences in stimulus levels; Gordon-Salant et al. used a stimulus level of 85 dB SPL, and this study used a stimulus level of 75 dB SPL. Goupell et al. (2017) obtained perceptual identification functions along the same DISH–DITCH continuum from older and younger adults using a level of 65 dB SPL, and the crossover points for the 65 dB SPL stimuli were later than those in our study. Differences in stimulus level appear to affect audibility and the ability to use duration cues for speech discrimination.

EEG

ABR

The click ABR was recorded to obtain a measure of peripheral function (possibly arising from cochlear synaptopathy) in addition to the audiometric thresholds. Equivalent Wave I amplitude and Wave V/I ratio between groups suggest that peripheral function did not play a role in group differences observed in perceptual performance (see Figure 8). Cochlear synaptopathy has been implicated in noise exposure sufficient to produce temporary threshold shift without permanent loss of outer hair cells in an animal model (Kujawa & Liberman, 2009), but the evidence of cochlear synaptopathy resulting from noise exposure in humans has been mixed. For example, veterans with high levels of noise exposure had reduced Wave I amplitudes compared with veterans with lower levels, despite normal outer hair cell function (Bramhall, Konrad-Martin, McMillan, & Griest, 2017). In contrast, young female adults with relatively higher levels of noise exposure history had equivalent Wave I amplitudes compared with young female adults with lower levels of noise exposure history (Prendergast et al., 2018). There is increasing evidence that synaptopathy may occur in the absence of excessive noise exposure in normal hearing aging mouse models (Sergeyenko et al., 2013). In humans with no history of otologic disease, postmortem temporal bone findings revealed evidence of cochlear synaptopathy (disruptions between synaptic ribbons and cochlear nerve terminals and cochlear nerve degeneration) without a significant loss of hair cells (Viana et al., 2015), suggesting that age-related cochlear synaptopathy findings in animal models may extend to humans. Although we did not find group differences, we noted considerable intersubject variability in the data, and it is possible that some individuals were affected by cochlear synaptopathy more than others. In addition, we note that our measure of Wave I amplitude is known to be variable (Bharadwaj, Masud, Mehraei, Verhulst, & Shinn-Cunningham, 2015; Bidelman, Pousson, Dugas, & Fehrenbach, 2018), and it is possible that a more rigorous measure that does not depend on amplitude would be more sensitive to cochlear synaptopathy.

FFR

The lack of aging effects on click ABR measures may suggest that age-related temporal processing deficits can arise from dysfunction in more central auditory regions, as has been found in previous studies investigating aging effects on the FFR (Anderson et al., 2012; Mamo et al., 2016; Presacco et al., 2015; Vander Werff & Burns, 2011). In particular, aging effects on FFR morphology have been demonstrated by Clinard and Cotter (2015), in which younger participants exhibited more accurate neural representations of stimulus waveforms than did older participants. Clinard and Cotter recorded FFRs to rising and falling frequency sweeps (starting/ending frequency of 400 Hz) at sweep rates ranging from 1333 to 6667 Hz. The older participants had lower STR correlations than the younger participants for all rising and falling sweep rates. Like the results of this study, such effects were observed even when the older participants had clinically normal hearing. Similarly, older rats have degraded responses compared with younger rats when presented with ramped and damped sinusoidally amplitude-modulated tones, which differ in their envelope shapes (Parthasarathy & Bartlett, 2011).

Reduced neural synchrony may underlie older participants' poorer neural representations of stimuli containing rapidly changing temporal and spectral properties, as in the case of speech stimuli (Anderson et al., 2012; Presacco et al., 2015, 2016). In contrast to the Anderson et al. (2012) and Presacco et al. (2016) studies, however, the current study only found aging effects in response to the temporal fine structure and not to the temporal envelope (see Figures 4 and 5). Ananthakrishnan, Krishnan, and Bartlett (2016), whose primary focus was to investigate the effects of age-related sensorineural hearing loss, found that older participants' disrupted phase locking affected the encoding of temporal fine structure more than the encoding of envelope cues. In the current study, reduction in temporal fine structure was equivalent in ONH and OHI participants compared with the YNH participants (see Figure 5), and the OHI participants in the study of Ananthakrishnan et al. (2016) were older than the ONH participants; therefore, aging effects may be a key factor in disrupted phase locking to the fine structure. To determine whether the lack of differences between the two older groups in our study was driven by aging effects, we repeated the analysis with the same restricted age groups as described in the discussion of the perceptual results and found a similar main effect of group on phase locking to the temporal fine structure, $F(2,38) = 6.47$, $p = .004$, $\eta_p^2 = .25$. The YNH group had higher PLF_{TFS} values than the ONH ($p = .024$) and OHI groups ($p = .012$), and the two older groups did not differ ($p = .997$). Furthermore, the lack of group difference on the envelope PLF remained, $F(2,38) = 0.25$, $p = .78$, $\eta_p^2 = .01$. Mamo et al. (2016) also found greater age-related reductions to the stimulus temporal fine structure than to the envelope; therefore, aging effects on temporal fine structure appear to be more pronounced than on the envelope.

We might attribute the absence of aging effects on envelope responses in the current study to the nature of the

speech stimuli we used. Most previous studies have used synthesized stimuli that contained sustained vocalic regions (Anderson et al., 2012; Mamo et al., 2016; Presacco et al., 2016), whereas this study examined phase locking to dynamic regions preceding the silent interval in naturally spoken words. The disruption to envelope encoding may be more pronounced for sustained stimuli. A previous study that evaluated aging effects on FFRs to synthesized speech syllables containing different vowel durations found that, compared with YNH participants, ONH participants had reduced amplitudes and phase locking for the stimulus containing the longer vowel (Presacco et al., 2015). Therefore, the envelope effect in previous studies may be due to inability to sustain neural synchrony.

Cortical

Prominent CAEP peaks correspond to different processes that occur prior to the listener's conscious percept of a stimulus. Earlier-occurring peaks P1 and N1 denote early preperceptual neural detection and triggering of attention to a presented stimulus, respectively (Näätänen & Winkler, 1999). Older participants increased N1 amplitudes suggest that greater neural resources are recruited for processing of the signal than for younger participants. It has been suggested that an age-related decrease in cortical connectivity may lead to redundant local processing of stimuli, resulting in overrepresentation or larger response amplitudes (Peelle, Troiani, Wingfield, & Grossman, 2010; Presacco et al., 2016). Another interpretation of these larger amplitudes derives from data collected from younger and older rats, showing an age-related loss of GABA markers in the superficial layer of the primary auditory cortex (Hughes et al., 2010), which might result in greater overall excitability. This higher excitability would lead to higher neuronal firing rates and larger response amplitudes in population responses. We also evaluated the later P2 and P1b peaks in response to our stimuli. P2 typically emerges at approximately 200 ms, possibly signifying auditory object identification of the stimulus (Näätänen & Winkler, 1999; Ross et al., 2013). The P1b peak corresponded to the final consonant of the stimuli. Prolonged P2 and P1b latencies may suggest that older participants require increased processing time to identify a stimulus because they have expended more neural resources encoding that stimulus compared with younger participants (see Figures 6 and 7). Maamor and Billings (2017) found aging effects on CAEP latency but not amplitude to 150-ms syllables /ba/ and /da/. Differences in patterns of results for amplitudes across the two studies may be attributed to the fact that Maamor and Billings (2017) used a longer interstimulus interval than was used in this study. Tremblay, Billings, and Rohila (2004) observed aging effects on both CAEP latency and amplitude, such that older participants exhibited prolonged latencies and exaggerated amplitudes compared with younger participants. Effects on amplitude, however, were only noted when complex stimuli (i.e., speech) were presented with faster rates but not slower presentation rates. Faster presentation rates may tax the older participants' processing at the level of the auditory

cortex more than for younger participants, requiring longer processing times and increased neural resources to identify complex stimuli, such as speech.

Group Differences Specific to DISH Versus DITCH

In the FFR, a Group $\times$ Word interaction was observed for phase locking to the temporal fine structure of the vowel /i/. This interaction was driven by an increase in the PLF in DITCH compared with DISH that was found only in the YNH group. The 60-ms silence duration in DITCH may have released backward masking effects of the fricative on the vowel that were present for DISH. Gehr and Sommers (1999) found aging effects on the time course of backward masking: Young participants obtained lower thresholds with masking delays greater than 6–8 ms, but older participants did not have lower thresholds with masking delays that were 20 ms (the longest delay used in the experiment).

Detection of the silent interval in DITCH may be analogous to a gap detection paradigm. It has been observed that aging influences listeners' ability to discriminate brief gap durations (Fitzgibbons & Gordon-Salant, 1994), which may be partially attributed to a reduction in the neurons in the IC that code shorter-duration gaps (Walton et al., 1998). The Walton et al. study also found delayed neural recovery in the IC in older mice compared with younger mice. Because the FFR has primary contributions from the IC, it might be expected that age-related degeneration in the ability to code brief gaps may specifically compromise representation of the word DITCH but not DISH, which does not contain a silent duration.

Similar group differences between words in the cortical responses were not observed. The latency delays may result from a nonspecific reduction in neural processing time.

Relationships Among Perceptual and EEG Variables

Midbrain variables (FFR waveform morphology and phase locking to temporal fine structure) and cortical variables (CAEP peak latency and amplitude) each contributed to variance in perceptual performance. Therefore, neural processing of the temporal characteristics of the stimulus appears to be a factor in the ability to discriminate between DISH and DITCH based on a silence duration cue. Previous studies have also demonstrated a link between electrophysiological and behavioral measures of temporal processing. Purcell, John, Schneider, and Picton (2004) recorded envelope-following responses (EFRs) to amplitude-modulated tones presented at rates varying from 20 to 600 Hz in younger and older participants. They observed aging effects on EFR amplitude only for frequencies above 100 Hz that have subcortical origins. They also found that the maximum detectable EFR frequency above the noise floor was correlated with the maximum perceptible modulation frequency and the behavioral gap detection threshold across participants. Other studies using speech stimuli compared the contributions of peripheral (audiometric thresholds) and midbrain representation of a speech stimulus (e.g.,

STR correlation and F1 magnitude) to the variance in speech-in-noise performance in older participants (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2013; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013). These studies found that the central (midbrain) variables made a significant contribution to the variance in self-assessment and/or sentence-in-noise recognition beyond what could be accounted for by pure-tone thresholds.

The above-mentioned studies found relationships among midbrain encoding and behavioral variables, but cortical processing also plays a key role in behavioral performance. Dimitrijevic et al. (2016) investigated aging effects on amplitude-modulated and EFR detection thresholds at varying modulation depths (2%–100%) at a rate of 41 Hz (arising from cortical sources) and found that older participants had reduced dynamic ranges compared with younger participants and that EFR thresholds were moderately correlated with behavioral thresholds across the two age groups. Using a gap detection experiment, Harris and Dubno (2017) found that cortical phase-locking values (4–8 Hz) and peak alpha frequency (7–14 Hz) related to lower gap detection thresholds, particularly in older participants. Finally, several studies have found relationships among cortical peak amplitudes/latencies and speech-in-noise performance measures in older adults (e.g., Billings, McMillan, Penman, & Gille, 2013; Billings, Penman, McMillan, & Ellis, 2015).

Studies investigating relationships among temporal processing or speech perception and putative measures of synaptopathy in humans have had mixed results. In young adults, the ABR Wave V slope (generated by latency shifts in decreasing signal-to-noise ratios) related to thresholds for envelope interaural time differences, a behavioral measure of binaural temporal processing (Mehraei et al., 2016). However, other studies have not found relationships among other abnormal ABR measures (Wave I amplitude or Wave V/I ratio) and perceptual performance (Grose et al., 2017; Guest et al., 2018). Although we found weak to moderate correlations between PTA and some of the midbrain and cortical variables, we did not find any correlations between Wave I amplitude or Wave V/I ratio and any of the behavioral or neural variables other than a weak relationship between Wave I amplitude and PTA ($r = -.34$). The findings in the current study suggest that both midbrain and cortical encoding contribute independently to perception of temporal speech contrasts.

The final model indicated that midbrain and cortical processing accounted for 39% of the variance in perceptual performance for the same DISH–DITCH stimuli. Although we screened our participants with measures designed to rule out cognitive dysfunction (the MoCA and the Wechsler Abbreviated Scale of Intelligence), we did not test these participants on other cognitive measures, such as speed of processing, working memory, or attention. Harris et al. (2012) found that attention effects on cortical responses were related to gap detection thresholds, suggesting that there were interacting effects of cognition and neural processing on perceptual performance. Future studies should consider

the inclusion of cognitive measures to provide a more comprehensive understanding of the factors contributing to speech understanding.

Conclusions

Age-related reductions in response morphology were observed in the FFR, and delayed cortical latencies were observed in the CAEP results. Peripheral measures (PTA or ABR) did not appear to contribute to perceptual performance. FFR and CAEP variables contributed significantly to variance in the ability to perceive the silence duration in the DISH–DITCH continuum. Therefore, degraded processing in midbrain and cortex of words containing silent durations, and perhaps in other temporal components, may contribute to degraded speech perception in older adults, particularly their ability to distinguish one word from another. Awareness and understanding of the specific mechanisms underlying older adults' temporal processing deficits is critical in the development of auditory training protocols and in the evaluation of the efficacy of those protocols. It is unknown whether these findings will generalize to other temporal components or to other populations of older adults. Future studies should extend this work to other temporal contrasts, such as vowel and transition duration.

Acknowledgments

This research was supported by the American Hearing Research Foundation (S. A.) and the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award R21DC015843 (S. A.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors wish to thank Calli Fodor and Erin Walter for their assistance with data collection and analysis.

References

- American National Standard Institute. (2010). *S3.6-2010, American National Standard specification for audiometers* (revision of ANSI S3.6-1996, 2004). New York, NY: Author.
- Ananthakrishnan, S., Krishnan, A., & Bartlett, E. (2016). Human frequency following response: Neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear and Hearing*, 37(2), e91–e103. <https://doi.org/10.1097/AUD.0000000000000247>
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience*, 32(41), 14156–14164. <https://doi.org/10.1523/jneurosci.2176-12.2012>
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2013). Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. *Journal of Speech, Language, and Hearing Research*, 56(1), 31–43. [https://doi.org/10.1044/1092-4388\(2012\)12-0043](https://doi.org/10.1044/1092-4388(2012)12-0043)
- Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing Research*, 300, 18–32. <https://doi.org/10.1016/j.heares.2013.03.006>

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300. Retrieved from <https://www.jstor.org/stable/2346101>
- Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *The Journal of Neuroscience*, 35(5), 2161–2172. <https://doi.org/10.1523/jneurosci.3915-14.2015>
- Bidelman, G. M., Pousson, M., Dugas, C., & Fehrenbach, A. (2018). Test–retest reliability of dual-recorded brainstem versus cortical auditory-evoked potentials to speech. *Journal of the American Academy of Audiology*, 29(2), 164–174. <https://doi.org/10.3766/jaaa.16167>
- Billings, C. J., McMillan, G. P., Penman, T. M., & Gille, S. M. (2013). Predicting perception in noise using cortical auditory evoked potentials. *Journal of the Association for Research in Otolaryngology*, 14(6), 891–903. <https://doi.org/10.1007/s10162-013-0415-y>
- Billings, C. J., Penman, T. M., McMillan, G. P., & Ellis, E. M. (2015). Electrophysiology and perception of speech in noise in older listeners: Effects of hearing impairment and age. *Ear and Hearing*, 36(6), 710–722. <https://doi.org/10.1097/aud.0000000000000191>
- Bramhall, N. F., Konrad-Martin, D., McMillan, G. P., & Griest, S. E. (2017). Auditory brainstem response altered in humans with noise exposure despite normal outer hair cell function. *Ear and Hearing*, 38(1), e1–e12. <https://doi.org/10.1097/aud.0000000000000370>
- Caspary, D. M., Ling, L., Turner, J. G., & Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *Journal of Experimental Biology*, 211(11), 1781–1791. <https://doi.org/10.1242/jeb.013581>
- Casseday, J. H., Ehrlich, D., & Covey, E. (2000). Neural measurement of sound duration: Control by excitatory–inhibitory interactions in the inferior colliculus. *Journal of Neurophysiology*, 84(3), 1475–1487. <https://doi.org/10.1152/jn.2000.84.3.1475>
- Ceponiene, R., Rinne, T., & Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clinical Neurophysiology*, 113(6), 870–882. [https://doi.org/10.1016/S1388-2457\(02\)00078-0](https://doi.org/10.1016/S1388-2457(02)00078-0)
- Ceponiene, R., Alku, P., Westerfield, M., Torki, M., & Townsend, J. (2005). ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology*, 42(4), 391–406. <https://doi.org/10.1111/j.1469-8986.2005.00305.x>
- Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research*, 323, 91–98. <https://doi.org/10.1016/j.heares.2015.02.002>
- Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, 7, 11070. <https://doi.org/10.1038/ncomms11070>
- de Cheveigne, A., & Simon, J. Z. (2008). Denoising based on spatial filtering. *Journal of Neuroscience Methods*, 171(2), 331–339. <https://doi.org/10.1016/j.jneumeth.2008.03.015>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dimitrijevic, A., Alsamri, J., John, M. S., Purcell, D., George, S., & Zeng, F. G. (2016). Human envelope following responses to amplitude modulation: Effects of aging and modulation depth. *Ear and Hearing*, 37(5), e322–335. <https://doi.org/10.1097/aud.0000000000000324>
- Dupuis, K., Pichora-Fuller, M. K., Chasteen, A. L., Marchuk, V., Singh, G., & Smith, S. L. (2015). Effects of hearing and vision impairments on the Montréal Cognitive Assessment. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 22(4), 413–437. <https://doi.org/10.1080/13825585.2014.968084>
- Eggermont, J. J. (2001). Between sound and perception: Reviewing the search for a neural code. *Hearing Research*, 157(1–2), 1–42. [https://doi.org/10.1016/S0378-5955\(01\)00259-3](https://doi.org/10.1016/S0378-5955(01)00259-3)
- Faure, P. A., Fremouw, T., Casseday, J. H., & Covey, E. (2003). Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. *The Journal of Neuroscience*, 23(7), 3052–3065. <https://doi.org/10.1523/JNEUROSCI.23-07-03052.2003>
- Fitzgibbons, P. J., & Gordon-Salant, S. (1994). Age effects on measures of auditory duration discrimination. *Journal of Speech and Hearing Research*, 37(3), 662–6670. <https://doi.org/10.1044/jshr.3703.662>
- Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, 6, 347. <https://doi.org/10.3389/fnagi.2014.00347>
- Galbraith, G. C., Arbagey, P. W., Branski, R., Comerchi, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *NeuroReport*, 6(17), 2363–2367.
- Gehr, S. E., & Sommers, M. S. (1999). Age differences in backward masking. *The Journal of the Acoustical Society of America*, 106(5), 2793–2799. <https://doi.org/10.1121/1.428104>
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517. <https://doi.org/10.1038/nn.3063>
- 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. *The Journal of the Acoustical Society of America*, 124(5), 3249–3260. <https://doi.org/10.1121/1.2982409>
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Barrett, J. (2006). Age-related differences in identification and discrimination of temporal cues in speech segments. *The Journal of the Acoustical Society of America*, 119(4), 2455–2466. <https://doi.org/10.1121/1.2171527>
- Goupell, M. J., Gaskins, C. R., Shader, M. J., Walter, E. P., Anderson, S., & Gordon-Salant, S. (2017). Age-related differences in the processing of temporal envelope and spectral cues in a speech segment. *Ear and Hearing*, 38(6), e335–e342. <https://doi.org/10.1097/aud.0000000000000447>
- Grose, J. H., Buss, E., & Hall, J. W., III. (2017). Loud music exposure and cochlear synaptopathy in young adults: Isolated auditory brainstem response effects but no perceptual consequences. *Trends in Hearing*, 21, 2331216517737417. <https://doi.org/10.1177/2331216517737417>
- Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing Research*, 364, 142–151. <https://doi.org/10.1016/j.heares.2018.03.008>
- Harris, K. C., & Dubno, J. R. (2017). Age-related deficits in auditory temporal processing: Unique contributions of neural dys-synchrony and slowed neuronal processing. *Neurobiology of Aging*, 53, 150–158. <https://doi.org/10.1016/j.neurobiolaging.2017.01.008>

- Harris, K. C., Wilson, S., Eckert, M. A., & Dubno, J. R. (2012). Human evoked cortical activity to silent gaps in noise: Effects of age, attention, and cortical processing speed. *Ear and Hearing*, 33(3), 330–339. <https://doi.org/10.1097/AUD.0b013e31823fb585>
- Heine, C., & Browning, C. J. (2002). Communication and psychosocial consequences of sensory loss in older adults: Overview and rehabilitation directions. *Disability and Rehabilitation*, 24(15), 763–773. <https://doi.org/10.1080/09638280210129162>
- Hughes, L. F., Turner, J. G., Parrish, J. L., & Caspary, D. M. (2010). Processing of broadband stimuli across A1 layers in young and aged rats. *Hearing Research*, 264(1–2), 79–85. <https://doi.org/10.1016/j.heares.2009.09.005>
- Humes, L. E., Kewley-Port, D., Fogerty, D., & Kinney, D. (2010). Measures of hearing threshold and temporal processing across the adult lifespan. *Hearing Research*, 264(1–2), 30–40. <https://doi.org/10.1016/j.heares.2009.09.010>
- Humes, L. E., Kidd, G. R., & Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Frontiers in Systems Neuroscience*, 7, 55. <https://doi.org/10.3389/fnsys.2013.00055>
- Jenkins, K. A., Fodor, C., Presacco, A., & Anderson, S. (2017). Effects of amplification on neural phase locking, amplitude, and latency to a speech syllable. *Ear and Hearing*, 39, 810–824. <https://doi.org/10.1097/aud.0000000000000538>
- Juarez-Salinas, D. L., Engle, J. R., Navarro, X. O., & Recanzone, G. H. (2010). Hierarchical and serial processing in the spatial auditory cortical pathway is degraded by natural aging. *The Journal of Neuroscience*, 30(44), 14795–14804. <https://doi.org/10.1523/jneurosci.3393-10.2010>
- Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss. *The Journal of Neuroscience*, 29(45), 14077–14085. <https://doi.org/10.1523/jneurosci.2845-09.2009>
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001–1010. <https://doi.org/10.1016/j.neuron.2007.06.004>
- Maamor, N., & Billings, C. J. (2017). Cortical signal-in-noise coding varies by noise type, signal-to-noise ratio, age, and hearing status. *Neuroscience Letters*, 636(Suppl. C), 258–264. <https://doi.org/10.1016/j.neulet.2016.11.020>
- Mamo, S. K., Grose, J. H., & Buss, E. (2016). Speech-evoked ABR: Effects of age and simulated neural temporal jitter. *Hearing Research*, 333, 201–209. <https://doi.org/10.1016/j.heares.2015.09.005>
- Martin, B. A., Tremblay, K., & Stapells, D. (2007). Principles and applications of cortical auditory evoked potentials. In R. F. Burkard, M. Don & J. J. Eggermont (Eds.), *Auditory evoked potentials. Basic principles and clinical application* (pp. 482–507). Baltimore: Lippincott, Williams & Wilkins.
- Mehraei, G., Hickox, A. E., Bharadwaj, H. M., Goldberg, H., Verhulst, S., Liberman, M. C., & Shinn-Cunningham, B. G. (2016). Auditory brainstem response latency in noise as a marker of cochlear synaptopathy. *The Journal of Neuroscience*, 36(13), 3755–3764. <https://doi.org/10.1523/jneurosci.4460-15.2016>
- Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and Clinical Neurophysiology*, 35(6), 665–667. [https://doi.org/10.1016/0013-4694\(73\)90223-X](https://doi.org/10.1016/0013-4694(73)90223-X)
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13(2), 201–233. <https://doi.org/10.1017/S0140525X00078407>
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125(6), 826–859. <https://doi.org/10.1037/0033-2909.125.6.826>
- Palmer, S. B., & Musiek, F. E. (2014). Electrophysiological gap detection thresholds: Effects of age and comparison with a behavioral measure. *Journal of the American Academy of Audiology*, 25(10), 999–1007. <https://doi.org/10.3766/jaaa.25.10.8>
- Parthasarathy, A., & Bartlett, E. L. (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>
- Peelle, J. E., Troiani, V., Wingfield, A., & Grossman, M. (2010). Neural processing during older adults’ comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cerebral Cortex*, 20(4), 773–782. <https://doi.org/10.1093/cercor/bhp142>
- Phillips, S. L., Gordon-Salant, S., Fitzgibbons, P. J., & Yeni-Komshian, G. (2000). Frequency and temporal resolution in elderly listeners with good and poor word recognition. *Journal of Speech, Language, and Hearing Research*, 43(1), 217–228. <https://doi.org/10.1044/jslhr.4301.217>
- Pichora-Fuller, M. K., Schneider, B. A., Benson, N. J., Hamstra, S. J., & Storzer, E. (2006). Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. *The Journal of the Acoustical Society of America*, 119(2), 1143–1155. <https://doi.org/10.1121/1.2149837>
- Prendergast, G., Tu, W., Guest, H., Millman, R. E., Kluk, K., Couth, S., . . . Plack, C. J. (2018). Supra-threshold auditory brainstem response amplitudes in humans: Test–retest reliability, electrode montage and noise exposure. *Hearing Research*, 364, 38–47. <https://doi.org/10.1016/j.heares.2018.04.002>
- Presacco, A., Jenkins, K., Lieberman, R., & Anderson, S. (2015). Effects of aging on the encoding of dynamic and static components of speech. *Ear and Hearing*, 36(6), e352–e363. <https://doi.org/10.1097/aud.0000000000000193>
- Presacco, A., Simon, J. Z., & Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of Neurophysiology*, 116(5), 2346–2355. <https://doi.org/10.1152/jn.00372.2016>
- Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America*, 116(6), 3581–3593. <https://doi.org/10.1121/1.1798354>
- Romero, S., Mañanas, M. A., & Barbanjo, M. J. (2006, August–September). *Quantitative evaluation of automatic ocular removal from simulated EEG signals: Regression vs. second order statistics methods*. Paper presented at the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY.
- Ross, B., Jamali, S., & Tremblay, K. L. (2013). Plasticity in neuro-magnetic cortical responses suggests enhanced auditory object representation. *BMC Neuroscience*, 14(1), 151. <https://doi.org/10.1186/1471-2202-14-151>
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, 22(15), 1417–1422. <https://doi.org/10.1016/j.cub.2012.05.025>
- Särelä, J., & Valpola, H. (2005). Denoising source separation. *Journal of Machine Learning Research*, 6(Mar), 233–272.
- Schatteman, T. A., Hughes, L. F., & Caspary, D. M. (2008). Aged-related loss of temporal processing: Altered responses to amplitude modulated tones in rat dorsal cochlear nucleus. *Neuroscience*, 154(1), 329–337. <https://doi.org/10.1016/j.neuroscience.2008.02.025>
- Schlägl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method

- of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, 118(1), 98–104. <https://doi.org/10.1016/j.clinph.2006.09.003>
- Schmiedt, R. A., Mills, J. H., & Boettcher, F. A.** (1996). Age-related loss of activity of auditory-nerve fibers. *Journal of Neurophysiology*, 76(4), 2799–2803. <https://doi.org/10.1152/jn.1996.76.4.2799>
- Schneider, B. A., & Hamstra, S. J.** (1999). Gap detection thresholds as a function of tonal duration for older and younger adults. *The Journal of the Acoustical Society of America*, 106, 371–380. <https://doi.org/10.1121/1.427062>
- Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., & Lamb, M.** (1994). Gap detection and the precedence effect in young and old adults. *The Journal of the Acoustical Society of America*, 95(2), 980–991. <https://doi.org/10.1121/1.408403>
- Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G.** (2013). Age-related cochlear synaptopathy: An early-onset contributor to auditory functional decline. *The Journal of Neuroscience*, 33(34), 13686–13694. <https://doi.org/10.1523/jneurosci.1783-13.2013>
- Sharma, A., Dorman, M., & Spahr, A.** (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, 23, 532–539. <https://doi.org/10.1097/01.AUD.0000042223.62381.01>
- Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R. J., & Näätänen, R.** (1998). Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. *Neuroscience Letters*, 251(2), 141–144. [https://doi.org/10.1016/S0304-3940\(98\)00529-1](https://doi.org/10.1016/S0304-3940(98)00529-1)
- Smith, J. C., Marsh, J. T., & Brown, W. S.** (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, 39(5), 465–472. [https://doi.org/10.1016/0013-4694\(75\)90047-4](https://doi.org/10.1016/0013-4694(75)90047-4)
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J.** (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience*, 16(13), 4240–4249. <https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996>
- Tremblay, K., Billings, C., & Rohila, N.** (2004). Speech evoked cortical potentials: Effects of age and stimulus presentation rate. *Journal of the American Academy of Audiology*, 15(3), 226–237. <https://doi.org/10.3766/jaaa.15.3.5>
- Tremblay, K., Piskosz, M., & Souza, P.** (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology*, 114, 1332–1343. [https://doi.org/10.1016/S1388-2457\(03\)00114-7](https://doi.org/10.1016/S1388-2457(03)00114-7)
- Vander Werff, K. R., & Burns, K. S.** (2011). Brain stem responses to speech in younger and older adults. *Ear and Hearing*, 32(2), 168–180. <https://doi.org/10.1097/AUD.0b013e3181f534b5>
- Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A., Santos, F., . . . Liberman, M. C.** (2015). Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue. *Hearing Research*, 327, 78–88. <https://doi.org/10.1016/j.heares.2015.04.014>
- Walton, J. P., Frisina, R. D., & O'Neill, W. E.** (1998). Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse. *The Journal of Neuroscience*, 18(7), 2764–2776. <https://doi.org/10.1523/JNEUROSCI.18-07-02764.1998>
- Working Group on Speech Understanding and Aging.** (1988). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83(3), 859–895. <https://doi.org/10.1121/1.395965>
- Zhu, J., & Garcia, E.** (1999). *The Wechsler Abbreviated Scale of Intelligence (WASI)*. New York, NY: The Psychological Corporation.