Objectives: Understanding speech in adverse listening environments is challenging for older adults. Individual differences in pure tone averages and working memory are known to be critical indicators of speech in noise comprehension. Recent studies have suggested that tracking of the speech envelope in cortical oscillations <8 Hz may be an important mechanism related to speech comprehension by segmenting speech into words and phrases (delta, 1 to 4 Hz) or phonemes and syllables (theta, 4 to 8 Hz). The purpose of this study was to investigate the extent to which individual differences in pure tone averages, working memory, and cortical tracking of the speech envelope relate to speech in noise comprehension in older adults.

Design: Cortical tracking of continuous speech was assessed using electroencephalography in older adults (60 to 80 years). Participants listened to speech in quiet and in the presence of noise (time-reversed speech) and answered comprehension questions. Participants completed Forward Digit Span and Backward Digit Span as measures of working memory, and pure tone averages were collected. An index of reduction in noise (RIN) was calculated by normalizing the difference between raw cortical tracking in quiet and in noise.

Results: Comprehension question performance was greater for speech in quiet than for speech in noise. The relationship between RIN and speech in noise comprehension was assessed while controlling for the effects of individual differences in pure tone averages and working memory. Delta band RIN correlated with speech in noise comprehension, while theta band RIN did not.

Conclusions: Cortical tracking by delta oscillations is robust to the effects of noise. These findings demonstrate that the magnitude of delta band RIN relates to individual differences in speech in noise comprehension in older adults. Delta band RIN may serve as a neural metric of speech in noise comprehension beyond the effects of pure tone averages and working memory.

Key words: Cortical oscillations, Older adults, Speech comprehension, Speech tracking.

INTRODUCTION

The ability to understand speech in noisy environments declines with age. Speech comprehension is essential to everyday life, whether understanding a conversation in a restaurant, near a construction zone, or over a noisy television. Older adults experience greater difficulty perceiving continuous speech in these challenging listening conditions relative to younger adults (CHABA 1988; Peelle et al. 2010; Smayda et al. 2016). Prior research suggests that although speech comprehension in quiet environments is preserved in older adults with normal hearing and in older adults with hearing loss, extraneous background noises adversely affect their speech comprehension (Pichora-Fuller 2003; Murphy et al. 2006). Behavioral studies have found that older adults perform worse on speech perception in noise tasks relative to younger adults (Pichora-Fuller et al. 1995; Tun et al. 2002; Smayda et al. 2016), although both age groups comprehend speech similarly in quiet conditions.

Multiple interactive factors can influence individual differences in speech in noise perception. Studies in older adults have demonstrated that peripheral hearing loss and deficits in cognitive functioning, moderated by motivation, are associated with impaired speech perception in noise performance (Humes & Roberts 1990; Pichora-Fuller et al. 1995; Fitzgibbons & Gordon-Salant 1996; Anderson et al. 2012; Humes 2013; Rönngberg et al. 2013; Peelle 2018). Temporal cues from the speech envelope are critical for speech perception (Shannon et al. 1995). Older adults often experience a decline in auditory temporal processing abilities, which can make speech perception difficult (CHABA 1988; Gordon-Salant & Fitzgibbons 1993; Gordon-Salant et al. 2011; Anderson et al. 2012). In addition to temporal processing deficits, aging is associated with elevated hearing thresholds, which can interfere with the ability to identify individual sounds that aid speech comprehension in challenging listening conditions (CHABA 1988; Humes & Roberts 1990; Fitzgibbons & Gordon-Salant 1996; Humes 2013). While speech perception can be affected by temporal processing deficits and hearing loss, speech comprehension also involves complex cognitive tasks that can be affected by deficits in cognitive functioning, even when hearing thresholds remain intact (Pichora-Fuller et al. 1995; Rönngberg et al. 2013). Working memory is a critical cognitive process for speech in noise comprehension, which allows the listener to attend to a target speaker while inhibiting irrelevant information in the form of background noise (Rönngberg et al. 2010; 2013; Füllgrabe & Rosen 2016; Dryden et al. 2017). Listeners utilize working memory to actively integrate words and phrases and access previously stored information to form a coherent representation of speech for comprehension (Pichora-Fuller et al. 1995). In combination, deficits in temporal processing, hearing status,
and working memory performance negatively impact speech in noise comprehension performance in older adults.

Prior studies evaluating the neurophysiological mechanisms underlying speech in noise comprehension in older adults used stimuli such as isolated words or single syllables (Wong et al. 2009, 2010). However, these stimuli are not representative of an everyday listening environment (Aiken & Picton 2008) and are not well-suited to assess real-world difficulties of speech in noise comprehension because they are rarely encountered in a natural setting (Hamilton & Huth 2018; Xie et al. 2019). Naturalistic stimuli, such as narrated passages and stories, provide an opportunity to examine speech and language processing encountered in everyday life (Aiken & Picton 2008; Moerel et al. 2012; Crosse et al. 2016; Huth et al. 2016; Broderick et al. 2018; Hamilton & Huth 2018; Decruy et al. 2019; Xie et al. 2019). Recently, studies have begun using natural language stimuli to evaluate the neural mechanisms underlying speech perception using neurophysiological methods like EEG (Horton et al. 2013; Di Liberto et al. 2015; Crosse et al. 2016; Petersen et al. 2016; Zoefel & VanRullen 2016; Dimitrijevic et al. 2017; Broderick et al. 2018; Das et al. 2018; Hambrock et al. 2018; Vanthornhout et al. 2018; Decruy et al. 2019; Destoky et al. 2019; Etard & Reichenbach 2019; Lesenfants et al. 2019) and magnetoencephalography (Ding & Simon 2012b; Ding & Simon 2013; Ding et al. 2014; Presacco et al. 2016b; Brodbeck et al. 2018; Destoky et al. 2019; Presacco et al. 2019; Vander Ghinste et al. 2019).

Cortical tracking of the temporal speech envelope has been shown to be a neural index of the encoding of dynamic temporal cues in natural speech that relates to speech in noise perception (Ding & Simon 2012b, 2013; Vanthornhout et al. 2018; Decruy et al. 2019). The most common approach adopted to evaluate cortical tracking of speech is estimating the linear mapping function between the temporal speech envelope and the neural response, and then assessing the fidelity of the neural processing in encoding the speech envelope (Lalor & Foxe 2010; Ding & Simon 2012b; Ding et al. 2014; Crosse et al. 2016; Di Liberto & Lalor 2017; Di Liberto et al. 2018a, 2018b; Vanthornhout et al. 2018; Decruy et al. 2019; Lesenfants et al. 2019; Xie et al. 2019). This measure of fidelity provides a direct index of how the dynamic acoustic cues of speech are encoded at the level of the cortex. Measures of cortical tracking of the speech envelope have been used to evaluate speech comprehension in the presence of competing talkers and background noise in normal hearing listeners (Ding & Simon 2012b; Columbic et al. 2013; Ding et al. 2014; Vanthornhout et al. 2018). This approach has specific advantages in disentangling the neural response to target speech from those of competing speech or noise. Cortical tracking of continuous speech in the presence of noise has been shown to relate to individual differences in speech reception thresholds in younger adults (Lesenfants et al. 2019). Thus, cortical tracking may provide an objective measure of speech perception in the presence of noise. Employing a similar approach to evaluate neural processing in older adults may aid in a better understanding of the mechanisms underlying individual differences in speech in noise comprehension.

Recent evidence implicates varying levels of contribution from different cortical oscillation frequency bands to the cortical tracking of the speech envelope. Low-frequency cortical oscillations (i.e., delta (1 to 4 Hz) and theta (4 to 8 Hz) bands) have distinct contributions to speech processing (Morillon et al. 2012; Peelle et al. 2013; Doelling et al. 2014; Ghizta 2017; Peffkou et al. 2017; Rimmle et al. 2018a). Specifically, theta oscillations align with the rhythm of phonemes and syllables in natural speech, occurring at a rate of 4 to 8 Hz (Ding et al. 2017). Stimulus-driven theta cortical tracking to syllabic rhythm has been attributed to be crucial for speech perception in quiet environments as well as in degraded listening situations (Howard & Poeppel 2010; Peelle et al. 2013; Doelling et al. 2014). While cortical tracking in theta may reflect syllable segmentation and speech clarity (Ding et al. 2014; Doelling et al. 2014; Molinaro & Lizarazu 2018; Etard & Reichenbach 2019), recent findings in younger adults suggests that delta band cortical tracking may aid in speech in noise comprehension (Molinaro & Lizarazu 2018; Etard & Reichenbach 2019). The rate that delta oscillations are suggested to parse speech may range from parsing syllables into words or words into phrases and other abstract linguistic structures (Ding et al. 2016). Prevailing research suggests delta oscillations may be important for speech segmentation based on prosodic cues and associated with combining segments of information for further cognitive-linguistic processing (Schroeder & Lakatos 2009; Ghizta 2017; Kösem & van Wassenhove 2017).

Speech analysis involves processing and decoding of hierarchical linguistic structures that occur on multiple timescales, such as phonemes, syllables, phrases, and sentences (Ding et al. 2016; Rimmle et al. 2018b). The role of the different cortical oscillations in speech processing has been explained using the analysis-by-synthesis account of speech perception (Ding & Simon 2014). Theta oscillations have been implicated in the speech analysis stage, which involves tracking sound onsets and parsing segmental cues, such as phonemes and syllables. In contrast, delta oscillations may aid in resynthesizing or filling-in missing segmental speech information by integrating cues spread across a longer timescale. This top-down modulatory feedback provided by delta oscillations is proposed to be important for speech perception in degraded listening situations (Schroeder & Lakatos 2009; Ghizta 2017; Kösem & van Wassenhove 2017).

Oscillations in alpha band (8 to 12 Hz) may aid selective attention and task switching in various speech intelligibility tasks (Strauß et al. 2014; Wilsch et al. 2015; McMahon et al. 2016; Dimitrijevic et al. 2017). Alpha oscillations play an important role in sensory processes involving working memory, which aids speech in noise comprehension (Bonnefond & Jensen 2012). Alpha oscillations have been shown to relate to listening effort for speech perception (Wüstmann et al. 2015; Dimitrijevic et al. 2019). However, alpha band’s contribution to tracking the speech envelope has not been extensively reported. An exact relationship between cortical tracking of speech by alpha oscillations and speech comprehension has yet to be established.

Older adults demonstrate a wide range of deficits in top-down and bottom-up processing of speech in the presence of noise and the underlying mechanisms are not completely understood (Anderson et al. 2010; Helfer et al. 2017; Maruthy et al. 2017). Few studies have examined cortical tracking of speech in older adults (Brodbeck et al. 2018; Decruy et al. 2019; Presacco et al. 2019), showing that cortical tracking of the speech envelope is greater in older adults than in younger adults (Brodbeck et al. 2018; Presacco et al. 2019). Increased cortical tracking of the speech envelope in older adults has been attributed to a greater reliance on acoustic cues to compensate for degraded input from the lower auditory centers (Brodbeck et al. 2018). Interestingly, the pattern of correlation between cortical tracking and speech comprehension has not been consistent. One study
demonstrated a significant correlation between cortical tracking and speech comprehension in older adults (Decruy et al. 2019), while another study did not observe such a pattern (Presacco et al. 2019). Additionally, it is plausible that individual differences in delta and theta band cortical tracking in older adults may obscure the relationship between cortical tracking and speech in noise comprehension. Understanding cortical tracking of speech in the presence of noise by delta, theta, and alpha bands in older adults will provide insights into mechanisms that may contribute to speech comprehension difficulties in adverse listening conditions.

In the present study, we investigated the extent to which cortical tracking of the speech envelope in the different cortical oscillation bands, as measured by EEG, may relate to individual differences in speech comprehension in the presence of noise in older adults. We evaluated cortical tracking of speech to an audiobook presented in quiet and in the presence of noise. We used noise consisting of time-reversed speech from a different segment of the same audiobook (i.e., same talker). Time-reversed speech is unintelligible (Arai 2010), but still contains partial phonetic information that is present in forward speech, even though lexical information is absent (Hoen et al. 2007). The time-reversed speech masker disrupts speech segmentation based on phonemic and syllabic boundaries. Thus, the time-reversed speech serves as a masker primarily interfering with the perception of segmental information, without causing semantic or syntactic interference. The listener must depend on building meaningful, linguistic units to segregate the target speech from the masker. This design is unlike prior studies in older adults where speech from different talkers were used for the target and maskers (Decruy et al. 2019; Presacco et al. 2019).

Using a time-reversed masker that was derived from the same talker as the target speech allowed us to evaluate cortical tracking and speech in noise comprehension without the interference of semantic and syntactic information from the masker. Therefore, cortical tracking of the speech envelope of a target speaker was assessed using a time-reversed speech masker to evaluate the mechanisms underlying individual differences in speech in noise comprehension in older adults that are not influenced by semantic and syntactic interferences. Considering the distinct contributions of the different cortical oscillation bands in cortical tracking of speech, we also examined the relationship between cortical tracking within the different cortical oscillation bands and speech in noise comprehension. We obtained pure tone averages (PTA) as a measure of hearing sensitivity and Digit Span scores as measures of working memory to use in a regression model to account for individual differences that may relate to speech in noise comprehension. In line with previous studies, we hypothesized that speech in noise comprehension would relate to delta band cortical tracking. We expect the direction of this relationship will shed light on the mechanisms underlying individual differences in speech in noise comprehension in older adults, beyond the effects of hearing sensitivity and working memory, as measured by PTA and Digit Span.

MATERIALS AND METHODS

Participants

Twenty-one older adults (11 females; ages 60 to 80 years, mean age = 71, SD = 4.80) were recruited from the Austin, Texas, community. Participants were a subset from a larger study investigating the effect of training on speech in noise perception in older adults who underwent 10-weeks of group piano classes, computer-based Listening and Communication Enhancement (LACE) auditory training, or no training. The current EEG study was conducted before the start of the training phase. All participants were native English speakers with no significant experience with a second language and reported no history of psychological or neurological disorders. Prior research suggests that speech processing in older adults is influenced by musical experience (Parbery-Clark et al. 2011). Therefore, we recruited participants who had less than 3 years of formal music training in their lifetime. After initial analyses, two participants were removed due to excessive noise in their raw EEG recordings that made the recordings inviable. Therefore, a total of 19 participants were used in the final analysis (9 females, mean age = 70.95 years, SD = 5.03). Each participant provided written, informed consent and received monetary compensation for their participation. This study was approved by the Institutional Review Board at The University of Texas at Austin.

Behavioral Measures

Pure Tone Audiometry • The hearing sensitivity for all participants was evaluated using pure tone audiometry. Air conduction hearing thresholds were measured via insert earphones for frequencies 250 through 8,000 Hz. Bone conduction hearing thresholds were measured for frequencies from 250 through 4,000 Hz. A modified Hughson and Westlake method (Carhart & Jerger 1959) was used to estimate the threshold at each frequency. None of the participants had an air-bone gap greater than 10 dB HL. Pure tone thresholds for all participants are shown in Figure 1. PTA was calculated by averaging the air conduction thresholds for 500, 1,000, and 2,000 Hz.

Working Memory • Working memory ability is known to affect speech comprehension (Pichora-Fuller et al. 1995; Ward et al. 2016). Therefore, participants completed the Forward Digit Span and Backward Digit Span tasks as measures of working memory. The tasks were administered by a trained researcher according to standard procedures as outlined by the Weschler Adult Intelligence Scale—Fourth Edition (WAIS-IV; Droidzick et al. 2012). In the Forward Digit Span task, participants were read a sequence of numbers and were asked to recall the numbers in the same order. For the Backward Digit Span task, participants were read a sequence of numbers and were asked to recall the numbers in reverse order. Each test consisted of eight sections with two sequences of digits per section. Sequences spanned two to nine digits in the Forward Digit Span and two to eight digits in the Backward Digit Span. Sequences increased in span size with each section. Testing was discontinued after accruing incorrect responses in both sequences of the same section. The digit span score represents the sum of sequences correctly identified across both tasks (Fig. 1C).

Stimuli

Participants listened to continuous speech from the public domain audiobook Alice’s Adventures in Wonderland (Carroll 1865). The story was read in American English by a male speaker and sampled at a frequency of 22.05 kHz. Long speaker pauses were manually truncated to a maximum of 500 ms, and the resulting speech from the audiobook was divided into segments approximately 60 sec in duration each
Segments began and ended with complete sentences. The mean syllable rate was ~6 Hz (Modulation spectrum shown in Supplemental Digital Content 1 [http://links.lww.com/EANDH/A687]). Participants listened to the story under three listening conditions: (1) quiet, (2) noise, and (3) time compressed. In the quiet condition, participants listened to the normal audiobook story with no background noise. The stimuli in the noise condition consisted of the same speech as in the quiet condition but overlapped with time-reversed speech (noise) from a different portion of the story unheard by the participants in this experiment. The target speech and noise began simultaneously and were presented at a signal-to-noise ratio of 0 dB. Pilot testing indicated 0 dB signal-to-noise ratio was optimal to prevent ceiling or floor effects in behavioral performance. The computer-based LACE intervention from the larger speech in noise perception study trains listeners for a variety of adverse listening conditions, including rapid speech perception. Therefore, a third EEG condition with time-compressed speech was included. Due to drastic differences in the speech envelope between the time-compressed condition stimuli and quiet condition stimuli, results from the time-compressed condition could not be directly compared with the quiet or noise conditions and are not reported in this study.

The order of conditions was counterbalanced across participants. Fifteen segments of the story were presented in each condition in chronological order, regardless of condition order, such that the storyline was preserved without repetitions or discontinuities. Participants were presented with two multiple-choice comprehension questions with four answer choices after each segment assessing speech comprehension of the preceding segment. Participants were provided feedback on their performance in the format of “You got X correct out of 2 in this trial.” Stimulus presentation was controlled by E-Prime 2.0.10 software (Schneider et al. 2002).

**EEG Acquisition**

Participants sat in an acoustically shielded booth. The stimuli were binaurally presented via insert earphones (ER-3A; Etymotic Research, Elk Grove Village, Illinois). Comfortable listening levels were individualized for each participant: Participants listened to three segments narrated by a male speaker and indicated changes to listening levels were established. Participants were instructed to listen to the story while maintaining visual fixation on a cross-hair centered on the screen in front of them, answer the comprehension questions to the best of their ability, and refrain from extraneous movement. Electrophysiological responses to continuous speech were amplified and digitized with BrainVision actiCHAMP amplifier and collected using BrainVision PyCorder 1.0.7 (Brain Products, Gilching, Germany) with 64-channel actiCAP active electrodes (Brain Products) secured in an elastic cap (EasyCap; [http://www.easycap.de]). Electrodes were placed on the scalp according to the International 10 to 20 system (Klem et al. 1999), and a common ground was placed at the AFz electrode site. Electrode impedance was less than 20 kΩ for all channels. Responses were recorded at a sampling rate of 25 kHz.

**EEG Preprocessing**

The EEG data were preprocessed using EEGLAB 14.1.2 (Delorme & Makeig 2004) in MATLAB 2019a (MathWorks Inc., Natick, Massachusetts) with the following steps. The raw EEG data were downsampled to 128 Hz to improve computational efficiency. The downsampled EEG was then filtered using minimum-phase causal windowed sinc finite impulse response filters. The high-pass filter cutoff frequency was set at 1 Hz with a filter order of 846, while the low-pass filter cutoff frequency was set at 15 Hz with a filter order of 212. Filter characteristics are shown in Supplemental Digital Content 2 ([http://links.lww.com/EANDH/A687]), based on recommendations of de Cheveigné and Nelken (2019). The filtered EEG was then re-referenced to the average of the two mastoid channels (O’Sullivan et al. 2014; Di Liberto et al. 2015; Di Liberto et al. 2018a). Re-referenced channels with electrical activity greater/lower
than 3 SDs of the surrounding channels were rejected, and their data were interpolated based on the activity in the surrounding channels using spherical spline interpolation. Artifacts in the EEG data were suppressed using artifact subspace reconstruction (ASR) (Mullen et al. 2015). Visually identified clean sections (~2 min) of the data were input as the calibration data for the ASR. The ASR cleaned data were then separated into epochs from −5 sec to 70 sec (re: trial onset), yielding 15 epochs in each condition. Independent component analysis was performed on the epoched data to remove eye-movement and muscle artifacts. Independent component analysis was performed using the infomax algorithm and was adjusted to extract only 50 components from the data to account for the reduced rank following channel interpolation and referencing. The independent components to be removed were visually identified based on time course, topography, and spectrum. Components corresponding to ocular and muscular activity were removed. The clean EEG was reconstructed from the remaining independent components.

For extracting the delta (lf = 1 Hz, order = 846; hf = 4 Hz, order = 424), theta (lf = 4 Hz, order = 424; hf = 8 Hz, order = 424), and alpha (lf = 8 Hz, order = 424, hf = 15 Hz, order = 424) bands from the EEG (Di Liberto et al. 2015), the cleaned data were further filtered using windowed sinc filters. This step avoided the confounds of having different artifact reduction parameters for the different bands. Filtering for the different EEG bands was done similarly to the earlier filtering by cascading the low- and high-pass filters. To avoid further accumulation of band specific delay, the group delay of the band pass filters was used to compensate for the delay introduced by these band pass filters.

**Estimation of Cortical Tracking to Speech**

Cortical tracking of the speech envelope was estimated using the Multivariate Temporal Response Function (mTRF) Toolbox in MATLAB (Crosse et al. 2016). We estimated the forward mapping of the target speech onto the EEG in both the quiet and noise conditions, which allowed us to assess the extent of neural representation of the target speech in both listening conditions. The EEG in the noise condition contains neural representation of both the target speech and the noise. The cortical tracking metric of the target speech in the noise condition quantified the fidelity of the representation of target speech in the EEG. Thus, the cortical tracking metric represents the tolerance of encoding the target speech in the presence of impeding effects of a masker. To exclude the possibility that our results were caused by the neural encoding of both the target speech and the masker, we ran an additional analysis that examined the extent of neural representation of the target speech while controlling for the neural representation of the masker, and our results did not differ (see Figure, Supplemental Digital Content 3, http://links.lww.com/EANDH/A687).

The multiband speech envelope was extracted using Hilbert decomposition of the output of 16 gamma tone filters equally spaced on the equivalent rectangular bandwidth scale (Slaney 1998) in the frequency range of 250 through 8,000 Hz. The amplitude of these multiband envelopes were raised to a power of 0.6 to mimic the compression by the inner ear (Vanthornhout et al. 2018; Decruy et al. 2019). These multiband envelopes were then downsampled from 22,500 Hz to 128 Hz to match the sampling rate of the EEG. Multivariate linear regression was used to derive the forward Temporal Response Function (TRF) \[ \text{TRF}(t, \text{chan}) = \sum_{i=0}^{15} \text{w}(t) \times \text{ENV}(t) \] between the stimulus envelope (ENV) and the EEG data (Crosse et al. 2016) in every channel and at different time lags. The TRF is a time series of the regression coefficients (beta weights) across different time lags. Higher beta weights indicate higher cortical tracking. The EEG epochs were trimmed based on the duration of the stimulus (~60 sec). The time lags used for TRF estimation were −100 to 450 ms.

A regularization parameter was applied during model estimation to enforce a smoothness constraint on the TRFs and reduce overfitting. The regularization parameter was optimized from 2° to 2° using cross-validation to obtain the best model fit across electrodes and participants (Crosse et al. 2016). The same regularization parameter was used for all subjects within each condition, with a different regularization parameter per condition based on model fit. The multiband TRFs were estimated using a 15-fold cross-validation, where 14 trials were used for TRF estimation in every iteration, that were then averaged across trials and used to predict the neural response in the left-out trial. The correlation between the predicted EEG and the actual EEG was estimated using Pearson’s correlation coefficient and served as a metric of cortical tracking (r-values). These steps were then iterated 15 times to estimate the r-values for all trials. The r-values were averaged across all 15 trials and served as the metric of similarity between the neural response and the stimulus envelope. r-values were estimated for each electrode. Based on the grand-averaged topographic map of r-values across participants and the different conditions, 18 fronto-central electrodes (FT8, FC6, FC4, FC2, T8, C6, C4, C2, FT7, FC5, FC3, FC1, T7, C5, C3, C1, FCz, Cz) were selected (Fig. 2). The averages of the r-values across these 18 fronto-central electrodes per participant and condition were used for further statistical analysis. The r-values were transformed using Fisher’s r-to-z transformation (hyperbolic arctangent) for statistical analysis. However, r-to-z has very minimal effects for correlations of magnitude less than 0.50, so the results of analyses reported in this study are identical if performed without the transformation.

**Statistical Analysis**

Scores on the comprehension questions were calculated based on the proportion of correct responses for each condition and were adjusted using the rationalized arcsine transformation (Studebaker 1985) to mitigate ceiling effects and normalize the distribution. Rationalized arcsine units (RAU) are interpretable like proportion correct except at extreme values (>0.94), which take on values between 100 and 123 (123 is equivalent to 100% correct) (Fig. 3). This RAU transformation was chosen to normalize the distribution of the comprehension scores for the linear models. The RAU comprehension scores were used for all of the following analyses.

We conducted paired-samples t-tests to compare raw cortical tracking between the quiet and noise conditions within each frequency band. The p-values were adjusted using a Bonferroni correction for the number of pairwise between-condition comparisons (4). To investigate the relationship between cortical tracking of the speech envelope and performance on the comprehension questions in the noise condition for delta, theta, alpha, and full bands, we calculated an index of reduction in noise (RIN). This RIN index takes the difference between a listener’s cortical tracking in
the quiet condition and in the noise condition and divides it by their cortical tracking in quiet to provide a normalized measure that reflects the amount of reduction in cortical tracking in noise, relative to the quiet condition: \( \frac{\text{Quiet} - \text{Noise}}{\text{Quiet}} \). A “large” RIN would indicate that raw cortical tracking in the noise condition was less than raw cortical tracking in the quiet condition. On the other hand, a “small” (or negative) RIN would suggest that cortical tracking of the speech envelope in the noise condition was similar to or greater than cortical tracking in the quiet condition.

**RESULTS**

As predicted, speech comprehension in the noise (\( M = 77.200, SD = 18.422 \)) condition was significantly lower than in the quiet condition (\( M = 92.455, SD = 11.316; t(18) = 3.563, p = 0.002 \)). When comparing raw cortical tracking (Fig. 4A) in quiet and noise, we observed a significant difference in the magnitude of raw theta band cortical tracking (Quiet: \( M = 0.118, SD = 0.058 \); Noise: \( M = 0.073, SD = 0.024; t(18) = 3.829, p = 0.005 \)) and raw full band cortical tracking (Quiet: \( M = 0.062, SD = 0.034 \); Noise: \( M = 0.042, SD = 0.017; t(18) = 3.475, p = 0.011 \)). However, we did not find a significant difference in the magnitude of raw delta band cortical tracking (Quiet: \( M = 0.054, SD = 0.037 \); Noise: \( M = 0.047, SD = 0.026; t(18) = 0.997, p = 1.000 \)) or raw alpha band cortical tracking (Quiet: \( M = 0.028, SD = 0.024 \); Noise: \( M = 0.016, SD = 0.015; t(18) = 2.555, p = 0.080 \)). A similar trend was observed in the TRF in the individual bands, where TRFs in full and theta bands showed higher beta weights for the quiet condition than the noise condition (Fig. 4B). The mean topographic maps of raw cortical tracking based on each electrode is also consistent with the differences obtained. The topographies of the full and theta bands were predominantly frontocentral, which is consistent with activity in the auditory cortices and the superior temporal plane (see Supplemental Digital Content 4, http://links.lww.com/EANDH/A687). Topography of the raw delta band cortical tracking showed a pattern consistent with activations of the cortical regions outside the core-auditory regions.

An initial analysis revealed that speech in noise comprehension and RIN within each frequency band were not significantly correlated (\( ps > 0.05 \)). Based on this null finding and the prior literature that suggests working memory and hearing status play a role in speech in noise comprehension (Pichora-Fuller et al. 1995, 2003; Murphy et al. 2006; Ward et al. 2016), the effects...
of working memory and PTA were regressed out from RIN and speech in noise comprehension scores. We fit separate linear models to RIN ($RIN_{band} \sim PTA_{better} + PTA_{poorer} + DigitSpan$) and speech in noise comprehension scores ($Comprehension_{Noise} \sim PTA_{better} + PTA_{poorer} + DigitSpan$), with PTA for each ear (separated by the better and poorer ear) and the combined Digit Span score (sum of Forward Digit Span and Backward Digit Span) as predictors in each linear model. The residuals of these models served as measures of RIN and speech in noise comprehension that were adjusted for individual differences in PTA and working memory. Using the residuals of these models, Spearman’s rank correlations were calculated to assess the relationship between the adjusted RIN within each frequency band and speech in noise comprehension when controlling for PTA and working memory. In the delta band, RIN and comprehension in the noise condition were significantly negatively correlated ($\rho = -0.656$, $p = 0.003$). However, no significant correlations were observed in theta ($\rho = 0.102$).
We also estimated linear models predicting comprehension in the noise condition including RIN as a predictor to assess the extent to which RIN significantly contributed to speech in noise comprehension by comparing the variance explained by each comprehension model (\(\text{Comprehension}_{\text{Noise}} - \text{PTA}_{\text{better}} + \text{PTA}_{\text{poorer}} + \text{DigitSpan} + \text{RIN}_{\text{band}}\)). The comprehension model including only PTA and Digit Span as predictors was not significant \((F(3,15) = 3.155, p = 0.056)\) with a Multiple \(R^2 = 0.387\) (Adj-\(R^2 = 0.264\)). Adding RIN in delta band as a fourth predictor to the comprehension model improved the significance \((F(4,14) = 6.019, p = 0.005)\) and increased the Multiple \(R^2 = 0.632\) (Adj-\(R^2 = 0.527\)). This addition of delta band RIN resulted in a significant improvement in model fit \((F(14,15) = 9.345, p = 0.009)\). We did not observe a significant change in model fit when RIN for theta, alpha, or full bands were added as a fourth predictor to the comprehension model (see Table 1 for model results). These findings suggest delta band RIN significantly increased the variance explained for comprehension scores in the noise condition.

Considering the age range of our older adult participants (60 to 80 years, \(M = 71, SD = 4.80\)), we estimated correlations between RIN and speech in noise comprehension when controlling for PTA, working memory, and age to examine the extent to which the age of our participants may have influenced our results. When age was added as a covariate, the significant negative correlation in the delta band was maintained \((\rho = −0.519, p = 0.024)\), as were the nonsignificant correlations in the full band \((\rho = 0.065, p = 0.792)\), theta band \((\rho = 0.295, p = 0.220)\), and alpha band \((\rho = −0.189, p = 0.436)\). Additionally, we compared the fit of the model for speech in noise comprehension that included PTA, working memory, and delta band RIN as predictors (reported above) and a model for speech in noise comprehension that included PTA, working memory, delta band RIN, and age as predictors. The addition of age as a predictor did not improve the model fit \((F(13,14) = 3.207, p = 0.097)\). These results suggest age was not a significant predictor in the correlation observed between delta band RIN and speech in noise comprehension within our sample of older adults.

### DISCUSSION

We examined the extent to which cortical tracking of the speech envelope relates to individual differences in speech in noise comprehension in older adults. The introduction of a masker resulted in a significant decrease in speech comprehension performance. Reduced hearing sensitivity (CHABA 1988; Pichora-Fuller et al. 1995; Fitzgibbons & Gordon-Salant 1996) and working memory (Pichora-Fuller et al. 1995; Rönnberg et al. 2013) are common factors affecting speech comprehension in older adults. We used a novel approach by assessing the relationship between RIN and speech in noise comprehension, beyond the effects of PTA and working memory. Unlike prior studies that investigated coherence between corresponding frequencies in the speech envelope and cortical oscillation bands (Howard & Poeppel 2010; Peelle & Davis 2012; Peelle et al. 2013; Doelling et al. 2014; Molinaro & Lizarazu 2018), the present study evaluated the distinct contributions of the different cortical oscillation bands in tracking the continuous speech envelope.

Raw cortical tracking of the speech envelope in theta band significantly reduced in the presence of noise (i.e., time-reversed speech), which is consistent with previous findings that the addition of noise reduces the salience of segmental cues of phonetic and syllabic boundaries (Howard & Poeppel 2010). The neural response in the noise condition represented the mixture of both the target speech and the masker. In other words, the cortical tracking to the target speech in the noise condition may also be influenced by cortical tracking to the noise itself. To this end, we performed an additional analysis to evaluate cortical tracking to the target speech in noise comprehension.

### TABLE 1. Linear Model: The Effect of Pure Tone Averages (PTA), Working Memory, and Cortical Tracking on Speech in Noise Comprehension Accuracy

| Band   | Model Predictors | Estimate | Standard Error | t value | p       |
|--------|------------------|----------|----------------|---------|---------|
| Full   | Intercept        | 117.425  | 22.244         | 5.279   | <0.001*** |
|        | PTA better ear   | −2.89    | 1.781          | −1.285  | 0.220   |
|        | PTA poorer ear   | 1.504    | 1.677          | 0.897   | 0.385   |
|        | Digit span       | −1.073   | 0.945          | −1.135  | 0.276   |
|        | Cortical tracking| −16.121  | 15.291         | −1.054  | 0.310   |
| Delta  | Intercept        | 95.280   | 18.930         | 5.033   | <0.001*** |
|        | PTA better ear   | −2.924   | 1.445          | −2.023  | 0.063   |
|        | PTA poorer ear   | 2.014    | 1.358          | 1.483   | 0.160   |
|        | Digit span       | −0.817   | 0.822          | −0.995  | 0.328   |
|        | Cortical tracking| −10.151  | 3.231          | −3.057  | 0.009*** |
| Theta  | Intercept        | 113.243  | 24.942         | 4.540   | <0.001*** |
|        | PTA better ear   | −2.458   | 1.961          | −1.254  | 0.230   |
|        | PTA poorer ear   | 1.636    | 1.832          | 0.893   | 0.387   |
|        | Digit span       | −1.086   | 1.013          | −1.072  | 0.302   |
|        | Cortical tracking| 2.890    | 16.200         | 0.178   | 0.861   |
| Alpha  | Intercept        | 113.909  | 22.271         | 5.115   | <0.001*** |
|        | PTA better ear   | −2.513   | 1.799          | −1.397  | 0.184   |
|        | PTA poorer ear   | 1.728    | 1.698          | 1.018   | 0.326   |
|        | Digit span       | −1.122   | 0.949          | −1.182  | 0.257   |
|        | Cortical tracking| −1.872   | 1.923          | −0.974  | 0.347   |

*Indicates significant predictors at alpha <0.01. **Indicates significant predictors at alpha <0.001.
speech while controlling for cortical tracking to the noise (see Figure, Supplemental Digital Content 3, http://links.lww.com/EANDH/A687), and our results did not change. Therefore, the magnitude of reduction in cortical tracking of the target speech in the presence of noise may indicate the neural resilience to the effects of noise.

As reflected by the reduction in raw theta band cortical tracking in the presence of noise, the phonetic ambiguity across competing speech streams and the associated energetic masking may impede segregation of the target speech from the masker, making the use of segmental boundaries for speech tracking and comprehension difficult. One possibility is that the reduced salience of phoneme and syllable boundaries due to competing speech streams may have led to a reduction in theta band cortical tracking. However, we were unable to find a relationship between theta band RIN and speech in noise comprehension, even after controlling for PTA and working memory. Thus, raw theta band cortical tracking of the speech envelope appears to be stimulus driven and does not directly relate to speech in noise comprehension. These findings suggest speech in noise comprehension may involve additional mechanisms beyond just phonemic and syllabic segmentation of speech alone. Delta band cortical tracking may be particularly important when the noise shares similar phonetic information as the target speech, which is the case with time-reversed speech derived from the same talker (Hoen et al. 2007). Therefore, listeners may utilize delta band cortical tracking for speech in noise comprehension, which may aid in parsing speech units based on linguistic structures spanning a longer timescale.

Alternatively, it could also be suggested that theta band cortical tracking in older adults may not be efficient in encoding and restoring speech information in the presence of a competing masker. Recent studies in younger adults have shown that theta band cortical tracking does not relate to speech comprehension in the presence of a competing masker (Molinaro & Lizarazu 2018; Etard & Reichenbach 2019). Few prior studies have shown that theta band cortical tracking correlates with speech perception (Ding & Simon 2013; Peelle et al. 2013; Doelling et al. 2014). However, these findings were reported for spectrally degraded speech without competing maskers. Thus, the converging evidence suggests that theta band cortical tracking of the speech envelope may have a limited role in aiding speech in noise comprehension in older adults, especially when the noise shares similar, but noninformational, segmental features.

In contrast to theta band, raw delta band cortical tracking did not significantly reduce in the presence of noise. This finding is consistent with prior research that delta band is robust to the effects of noise, while theta band reduces substantially with the introduction of background noise (Ding & Simon 2014; Ding et al. 2014). The magnitude of delta band cortical tracking may depend on the amount of information the listener can construct from an impoverished acoustic scene by using higher-order linguistic cues to segregate a target speaker from background noise. Our results in older adults demonstrate that the effect of noise on the robustness of raw delta band cortical tracking may be individual-specific and relates strongly with speech in noise comprehension. This result is similar to previous findings where delta band cortical tracking correlated with speech intelligibility in the presence of noise in younger adults (Ding & Simon 2013; Doelling et al. 2014) and was robust to the disruptive effects of noise.

Interestingly, we observed a significant, negative correlation between speech in noise comprehension and delta band RIN, even though we did not observe a significant difference between raw delta band cortical tracking in quiet and in noise. Prior studies in younger adults have observed a positive relationship between raw delta band cortical tracking and speech comprehension (Ding & Simon 2013; Doelling et al. 2014). Lower RIN values suggest that the cortical tracking was robust to the effects of noise, while higher RIN suggests that cortical tracking decreased in the presence of noise. We specifically focused on the reduction in cortical tracking between the quiet and noise conditions, relative to cortical tracking in the quiet condition, to better understand the extent to which the change in magnitude of raw cortical tracking between listening conditions relates to individual differences in speech in noise comprehension, rather than directly examining raw cortical tracking and its relationship with speech in noise comprehension. After controlling for the effects of PTA and working memory, listeners with a greater delta band RIN tended to have lower speech in noise comprehension scores, while listeners with a smaller delta band RIN tended to have higher speech in noise comprehension scores. Our index of RIN is a novel approach, and additional studies are needed to extend these results. While we did control for PTA in the current study, future studies could explore presenting stimuli at individual participants’ frequency-specific sensation level as a better method for controlling for the effects of audibility (McClannahan et al. 2019). We presented continuous speech masked by time-reversed speech in the current study. Prior research has demonstrated that cortical tracking of speech in older adults is less degraded when the noise is unintelligible (Presacco et al. 2016a). Time-reversed speech is unintelligible, and the acoustic envelope differs from the acoustic envelope of forward speech. Further studies are needed to examine the extent to which the relationship between delta band RIN and speech in noise comprehension holds for other noise maskers, including forward speech and energetic noise.

Prior connectivity analyses using magnetoencephalography suggest that top-down modulation of auditory cortices by the inferior frontal gyrus during speech perception may be mediated by delta oscillations (Park et al. 2015). The topography of raw delta band cortical tracking (Fig. 4C) in the current study suggested activation of cortical regions beyond the auditory regions (qualitatively based on the topography differing from the canonical N1 wave, which is primarily generated along the auditory regions in the brain; see Figure, Supplemental Digital Content 4, http://links.lww.com/EANDH/A687), aligning with other prior findings that older adults recruit additional cortical regions when listening to speech (Wong et al. 2009; Brodbeck et al. 2018). Delta oscillations may have a multifaceted role in structure-building and top-down modulation of cortical encoding that could be vital for speech in noise comprehension. This finding aligns with the analysis-by-synthesis account of speech comprehension that suggests structure-building (top-down modulation) mediated by cortical tracking in delta band, may modulate and improve overall encoding of the acoustic envelope (Ding & Simon 2014). Therefore, delta band cortical tracking might aid in online recalibration of the speech encoding mechanism, resulting in enhanced speech in noise comprehension. The reader is cautioned about the interpretation of cortical sources. The inferences made here regarding the cortical sources of speech tracking that were based on the topographies
are qualitative, and are not based on source localization metrics, as we did not have structural MRI scans from our participants.

Prior work in younger adults has indicated that alpha oscillatory power may play an inhibitory role in speech processing in both quiet and challenging listening environments, relating to cognitive processes such as working memory and attention (Strauß et al. 2014; Dimitrijevic et al. 2017; Deng et al. 2019, 2020). The previous studies examined overall changes in alpha oscillatory power to study the effects of challenging listening environments. In contrast to alpha oscillatory power, cortical tracking examines the extent to which alpha oscillations phase lock to the stimulus envelope. These two metrics shed light on different processes, which mark an important distinction between the current study and prior research. Therefore, we posit that the role of alpha oscillations in cortical tracking of the speech envelope may be limited, suggesting that alpha oscillations may serve a different but important role beyond tracking the stimulus envelope in aiding speech in noise comprehension.

The present study focused on the relationship between cortical tracking of continuous speech and speech in noise comprehension in older adults. While our results did not differ when controlling for age in addition to PTA and working memory, further studies are needed to extend the current findings to younger adults and the extent to which our findings are because of aging. Additionally, we acknowledge that the comprehension task used here has not been validated in reference to other standardized metrics of speech perception. However, the focus of the current study was to evaluate the relationship between cortical tracking and in-task speech perception. Several other cortical tracking studies have used similar experimental paradigms that had participants answer questions about continuous speech stimuli obtained from audiobooks (Ding & Simon 2012a, 2013; Broderick et al. 2018). Attempts should be made to standardize the comprehension metrics for continuous speech tasks used in cortical tracking studies, which will help to compare findings across studies.

In conclusion, our results indicate a direct relationship between delta band RIN (as measured by cortical tracking) and individual differences in older adults’ speech in noise comprehension, beyond the effects of PTA and working memory. While aging is generally accompanied with a decline in pure tone thresholds and working memory (CHABA 1988; Humes & Roberts 1990; Pichora-Fuller et al. 1995, 2003; Humes 2013), pure tone thresholds and working memory performance do not provide a full picture of speech processing abilities in an everyday listening environment. After thorough experimental evaluation and validation, cortical tracking of the speech envelope in noise could potentially be used as a clinical measure of speech perception. Considering the converging evidence that delta band cortical tracking may be an indicator of speech in noise comprehension, the findings from the present study encourage further investigation of speech processing abilities using cortical tracking of the speech envelope to continuous, naturalistic speech.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health grants (1R01DC013315 and 1R01DC015504) to B. C., the National Institute on Aging National Research Service Award (1F31AG052308-01A1) to K. E. S., and the National Institutes of Health grant (T32-DC011499) to K. Kandler and B. Yates (trainees: J. McHaney). This work was also supported in part by the resources provided through the University of Pittsburgh Center for Research Computing (https://doi.org/10.1002/open.201700046). J. R. M designed and performed experiment, analyzed data, and wrote the paper; G. N. G analyzed data and wrote the paper; K. E. S. designed experiment and collected data; B. D. Z. analyzed data and provided statistical analysis; and B. C. designed experiment, provided statistical analysis, and provided critical revision. J. R. M and G. N. G. are co-first authors.

The authors have no conflicts of interest, financial, or otherwise to declare.

Address for correspondence: Bharath Chandrasekaran, University of Pittsburgh, Department of Communication Science and Disorders, 6035 Forbes Tower, Atwood and Sennott Sts, Pittsburgh, PA 15260, USA. E-mail: b.chandra@pitt.edu

Received October 16, 2019; accepted June 14, 2020.

REFERENCES

Aiken, S. J., & Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear Hear*, 29, 139–157.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., et al. (2012). Aging affects neural precision of speech encoding. *J Neurosci*, 32, 14156–14164.

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

Arai, T. (2010). Masking speech with its time-reversed signal. *Acoust Sci Technol*, 31, 188–190.

Bonnefon, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distractors. *Curr Biol*, 22, 1969–1974.

Brodebeck, C., Presacco, A., Anderson, S., et al. (2018). Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust*, 104, 774–777.

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., et al. (2018). Electro-physiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Curr Biol*, 28, 803–809.

Carhart, R., & Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *J Speech Hear Res*, 24, 330–345.

Carroll, L. (1865). *Alice’s adventures in wonderland*. Retrieved from https://librivox.org/alices-adventures-in-wonderland-by-lewis-carroll-5.

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics). (1988). Speech understanding and aging. *J Acoust Soc Am*, 83, 859–895.

de Cheveigné, A., & Nelken, I. (2019). Filters: When, why, and how (not) to use them. *Neuron*, 102, 280–293.

Crosby, M. J., Di Liberto, G. M., Bednar, A., et al. (2016). The Multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Front Hum Neurosci*, 10, 604.

Das, N., Bertrand, A., Francart, T. (2018). EEG-based auditory attention detection: Boundary conditions for background noise and speaker positions. *J Neural Eng*, 15, 066017.

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.

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.

Deng, Y., Choi, I., Shinn-Cunningham, B. (2020). Topographic specificity of alpha power during auditory spatial attention. *Neuroimage*, 207, 116360.

Deng, Y., Reinhart, R. M., Choi, I., et al. (2019). Causal links between parietal alpha activity and spatial auditory attention. *Elife*, 8.

Destoy, F., Philippe, M., Bertels, J., et al. (2019). Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. *Neuroimage*, 184, 201–213.

Di Liberto, G. M., Crosse, M. J., Lalor, E. C. (2018a). Cortical measures of phoneme-level speech encoding correlate with the Perceived Clarity of Natural Speech. *eNeuro*, 5.

Di Liberto, G. M., & Lalor, E. C. (2017). Indexing cortical entrainment to natural speech at the phonemic level: Methodological considerations for applied research. *Hear Res*, 348, 70–77.

Di Liberto, G. M., Lalor, E. C., Millman, R. E. (2018). Causal cortical dynamics of a predictive enhancement of speech intelligibility. *Neuroimage*, 166, 247–258.

Di Liberto, G. M., O’Sullivan, J. A., Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Curr Biol*, 25, 2457–2465.
Humes, L. E. (2013). Understanding the speech-understanding problems of older adults. *Am J Audiol*, 22, 303–305.

Humes, L. E., & Roberts, L. (1990). Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility. *J Speech Hear Res*, 33, 726–735.

Huth, A. G., de Heer, W. A., Griffiths, T. L., et al. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532, 453.

Klem, G., Luders, H., Jasper, R., et al. (1999). The ten-twenty electrode system of the International Federation. *Electroencephalography Clin Neurophysiol*, 52, 3–6.

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Lang Cogn Neurosci*, 32, 536–544.

Lalor, E. C., & Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *J Neurosci*, 31, 189–193.

Lesenfants, D., Vanthornhout, J., Verschueren, E., et al. (2019). Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. *Hear Res*, 380, 1–9.

Maruthy, S., Kumar, A. U., Ganateja, G. N. (2017). Functional interplay between the putative measures of rostral and caudal efferent regulation of speech perception in noise. *J Assoc Res Otolaryngol*, 18, 635–648.

McClannahan, K. S., Backer, K. C., Tremblay, K. L. (2019). Auditory evoked responses in older adults with normal hearing, untreated, and treated age-related hearing loss. *Ear Hear*, 40, 1106–1116.

McMahon, C. M., Boisvert, L., de Lissa, P., et al. (2016). Monitoring alpha oscillations and pupil dilation across a performance-intensity function. *Front Psychol*, 7, 1–12.

Morel, M., De Martino, F., Formisano, E. (2012). Processing of natural sounds in human auditory cortex: Tonotopy, spectral tuning, and relation to voice sensitivity. *J Neurosci*, 32, 14205–14216.

Molinaro, N., & Lizarazu, M. (2018). Delta (but not theta)-band cortical entrainment involves speech-specific processing. *J Neurosci*, 48, 2642–2650.

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., et al. (2012). Asymmetric function of theta and gamma activity in syllable processing: An intra-cortical study. *Front Psychol*, 3, 1–9.

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., et al. (2015). Real-time neuro-imaging and cognitive monitoring using wearable dry EEG. *IEEE Trans Biomed Eng*, 62, 2553–2567.

Murphy, D. R., Daneman, M., Schneider, B. A. (2006). Why do older adults have difficulty following conversations? *Psychol Aging*, 21, 49–61.

O’Sullivan, J. A., Power, A. J., Mesgarani, N., et al. (2014). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb Cortex*, 25, 1697–1706.

Parbery-Clark, A., Strait, D. L., Anderson, S., et al. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *J Speech Lang Hear Res*, 54, 1391–1406.

Park, H., Ince, R. A. A., Schyns, P. G., et al. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr Biol*, 25, 1649–1653.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear Hear*, 39, 204–214.

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front Psychol*, 3, 1–33.

Peelle, J. E., Gross, J., Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb Cortex*, 23, 1378–1387.

Peelle, J. E., Troiani, V., Wingfield, A., et al. (2010). Neural processing during older adults’ comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cereb Cortex*, 20, 773–782.

Pekouz, M., Arnal, L. H., Fontolan, L., et al. (2017). Θ-Band and B-band neural activity reflects independent syllable tracking and comprehension of time-compressed speech. *J Neurosci*, 37, 7930–7938.

Petersen, E. B., Wüstmann, M., Obleser, J., et al. (2016). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J Neuropsychiatry Clin Neurosci*, 28, 14205–14216.

Presacco, A., Simon, J. Z., Anderson, S. (2016a). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neuropsychiatry Clin Neurosci*, 28, 14205–14216.
Presacco, A., Simon, J. Z., Anderson, S. (2016b). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol*, 116, 2346–2355.

Presacco, A., Simon, J. Z., Anderson, S. (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS One*, 14, e0213899.

Rimmele, J.M., Gross, J., Molholm, S., et al. (2018a). Brain oscillations in human communication. *Front Hum Neurosci*, 12, 39.

Rimmele, J. M., Morillon, B., Poeppel, D., et al. (2018b). Proactive sensing of periodic and aperiodic auditory patterns. *Trends Cogn Sci*, 22, 870–882.

Rönnberg, J., Lunner, T., Zekveld, A. A., et al. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front Syst Neurosci*, 7, 1–17.

Rönnberg, J., Rudner, M., Lunner, T., et al. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise Heal*, 12, 263.

Schneider, W., Eschman, A., Zuccolotto, A. (2002). E-Prime: User’s guide. Psychology Software Incorporated.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci*, 32, 9–18.

Shannon, R. V., Zeng, F. G., Kanath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.

Slaney, M. (1998). Auditory toolbox. *Interval Res Corp Tech Rep*, 10.

Smayda, K. E., Van Engen, K. J., Maddox, W. T., Chandrasekaran, B. (2016). Audio-visual and meaningful semantic context enhancements in older and younger adults. *PLoS One*, 11, e0152773.

Strauß, A., Wöstmann, M., Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci*, 8, 1–7.

Studebaker, G. A. (1985). A “rationalized” arcsine transform. *J Speech Hear Res*, 28, 455–462.

Tun, P. A., O’Kane, G., Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychol Aging*, 17, 453–467.

Vander Ghinast, M., Bourguignon, M., Niesen, M., et al. (2019). Cortical tracking of speech-in-noise develops from childhood to adulthood. *J Neurosci*, 39, 2938–2950.

Vanthornhout, J., Decuyr, L., Wouters, J., et al. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *J Assoc Res Otolaryngol*, 19, 181–191.

Ward, C. M., Rogers, C. S., Van Engen, K. J., et al. (2016). Effects of age, acoustic challenge, and verbal working memory on recall of narrative speech. *Exp Aging Res*, 42, 97–111.

Wilsch, A., Henry, M. J., Herrmann, B., et al. (2015). Alpha oscillatory dynamics index temporal expectation benefits in working m. *Cereb Cortex*, 25, 1938–1946.

Wong, P. C., Ettlinger, M., Sheppard, J. P., et al. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear Hear*, 31, 471–479.

Wong, P. C., Jin, J. X., Gunasekera, G. M., et al. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, 47, 693–703.

Wöstmann, M., Herrmann, B., Wilsch, A., Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J Neurosci*, 35, 1458–1467.

Xie, Z., Reetzke, R., Chandrasekaran, B. (2019). Machine learning approaches to analyze speech-evoked neurophysiological responses. *J Speech Lang Hear Res*, 62, 587–601.

Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *Neuroimage*, 124(Pt A), 16–23.