

RESEARCH ARTICLE

Sensory Processing

Effects of aging on cortical representations of continuous speech

 I. M. Dushyanthi Karunathilake,¹ Jason L. Dunlap,² Janani Perera,²  Alessandro Presacco,³
 Lien Decruy,³  Samira Anderson,²  Stefanie E. Kuchinsky,⁴ and  Jonathan Z. Simon^{1,3,5}

¹Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland, United States;

²Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland, United States; ³Institute for Systems Research, University of Maryland, College Park, Maryland, United States; ⁴Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, Maryland, United States; and ⁵Department of Biology, University of Maryland, College Park, Maryland, United States

Abstract

Understanding speech in a noisy environment is crucial in day-to-day interactions and yet becomes more challenging with age, even for healthy aging. Age-related changes in the neural mechanisms that enable speech-in-noise listening have been investigated previously; however, the extent to which age affects the timing and fidelity of encoding of target and interfering speech streams is not well understood. Using magnetoencephalography (MEG), we investigated how continuous speech is represented in auditory cortex in the presence of interfering speech in younger and older adults. Cortical representations were obtained from neural responses that time-locked to the speech envelopes with speech envelope reconstruction and temporal response functions (TRFs). TRFs showed three prominent peaks corresponding to auditory cortical processing stages: early (~50 ms), middle (~100 ms), and late (~200 ms). Older adults showed exaggerated speech envelope representations compared with younger adults. Temporal analysis revealed both that the age-related exaggeration starts as early as ~50 ms and that older adults needed a substantially longer integration time window to achieve their better reconstruction of the speech envelope. As expected, with increased speech masking envelope reconstruction for the attended talker decreased and all three TRF peaks were delayed, with aging contributing additionally to the reduction. Interestingly, for older adults the late peak was delayed, suggesting that this late peak may receive contributions from multiple sources. Together these results suggest that there are several mechanisms at play compensating for age-related temporal processing deficits at several stages but which are not able to fully reestablish unimpaired speech perception.

NEW & NOTEWORTHY We observed age-related changes in cortical temporal processing of continuous speech that may be related to older adults' difficulty in understanding speech in noise. These changes occur in both timing and strength of the speech representations at different cortical processing stages and depend on both noise condition and selective attention. Critically, their dependence on noise condition changes dramatically among the early, middle, and late cortical processing stages, underscoring how aging differentially affects these stages.

auditory aging; speech in noise; speech processing; stimulus reconstruction; temporal processing

INTRODUCTION

Speech communication is crucial in day-to-day interactions, and our interactions with others depend heavily on our ability to understand speech in a variety of listening conditions. Speech comprehension becomes increasingly difficult in a noisy environment and, critically, degrades further with aging. Compared with younger adults, older adults

have been observed to exhibit greater difficulty in auditory tasks in the presence of background noise, whether in relatively simple paradigms such as pitch discrimination (1) and gap detection (2) or in acoustically complex paradigms such as speech listening in noise (3, 4). Because poor speech comprehension in noise is associated with adverse psycho-social effects (5, 6), depression (7), and dementia (8), identifying age-related changes in the neural mechanisms that underlie



speech-in-noise difficulties may be critical for developing remediations that improve communication and quality of life among older adults. The present study aimed to investigate the timing and fidelity of cortical auditory processing in a more ecologically valid naturalistic speech paradigm that could be incorporated into both diagnostic evaluations and treatments aimed at speech comprehension problems in older adults.

Numerous studies have demonstrated important age-related anatomical and functional changes in auditory cortical pathways that may contribute to listening difficulties. In studies with human subjects, exaggerated neural activity in the auditory cortex has been observed in older adults during speech and speech-in-noise tasks (9–14), even when age-related hearing loss is mild, potentially reflecting changes in the central auditory system that could contribute to speech comprehension difficulties in older adults. Animal studies have also reported age-related excitatory and inhibitory imbalance in auditory cortex (15, 16), leading to altered neural coding.

Age-related changes in auditory, linguistic, and cognitive processes can all influence speech understanding (for a review, see Ref. 17). Age-related changes in cognitive functions limit speech understanding among older adults (18), in addition to other combinations of anatomical, functional, and cognitive factors (19). Therefore, audiometric measurements of the peripheral auditory system alone may not be sufficient to evaluate and manage the speech understanding difficulties reported by older adults.

The human neurophysiology underlying age-related changes in the timing and fidelity of encoding of connected speech is not well understood. High-temporal resolution recording methods such as magnetoencephalography (MEG) and electroencephalography (EEG) are well suited to accurately estimate such neural responses and the effects of aging on them. Much of this research routinely focuses on auditory evoked responses that quantify the central auditory functions obtained by averaging response to many repetitions of simple stimuli. Such studies often find that the cortical auditory event potential (CAEP), e.g., the P1-N1-P2 complex, is enhanced with aging (20–23). Reduced inhibitory neurotransmitters in the ascending auditory pathway may contribute to these enhanced neural responses (24). In addition, these studies have also reported delayed N1 and P2 peaks in older adults compared with younger adults (for a review, see Ref. 25). Enhanced and delayed neural responses may demonstrate at least partial compensation for and restoration of degraded subcortical input at the cortex.

However, simple stimuli (e.g., tones, clicks, and single speech syllables) do not well replicate real-world listening, where the ultimate goal is speech comprehension (26). Computational tools have been developed that can analyze neural responses to continuous speech, typically in terms of speech encoding and decoding models. Speech envelope reconstruction analysis and temporal response function (TRF) analysis, respectively, measure neural speech processing as linear decoding and encoding methods and can be used for both attended (foreground) and unattended (background) speech streams. In contrast, CAEP responses cannot straightforwardly segregate the neural responses to individual acoustic components in the stimulus (e.g., latency effects

of selective attention). In this study, we extend the limitations of CAEP studies with a more naturalistic speech paradigm and by investigating the timing and fidelity of both attended and unattended speech responses that could explain cortical processing differences in older adults.

In recent years, both EEG and MEG studies analyzing the reconstruction accuracy of the speech envelope have reported an enhanced cortical representation of the attended speech envelope in older adults. Reconstruction analysis integrates over a long time window (typically 500 ms), however, making estimation of the latency time course of age-related overrepresentation more difficult. Presacco et al. (14) showed that older adults' cortical ability to track the speech envelope is significantly reduced when decreasing the integration window to 150 ms, indicating that, at least, longer latencies contribute significantly. In this study, we replicate and extend the temporal analysis window results of Presacco et al. (14), using a nonlinear mixed-effects modeling approach [i.e., generalized additive mixed-effects models (GAMMs)] to better understand the time course of speech processing.

Although temporal integration window analysis is robust because of its integration over sensors and time, detailed temporal information can be gained directly from a TRF analysis (27, 28). Prominent peaks in the TRF, the $M50_{TRF}$, $M100_{TRF}$, and $M200_{TRF}$, can be ascribed to different auditory processing stages in the cortex with the corresponding latencies (29). In analogy to the P1-N1-P2 CAEP peaks at the corresponding latencies, it has been suggested that the early $M50_{TRF}$ (P1 like) peak dominantly reflects the low-level acoustic features such as changes in acoustic power (30, 31), whereas the $M100_{TRF}$ (N1 like) peak reflects processing of selectively attended features (31). Similarly, the $M50_{TRF}$ has been shown to depend more on the properties of the acoustic stimulus than the focus of selective attention, whereas the $M100_{TRF}$ shows the opposite (32, 33). The late $M200_{TRF}$ (P2 like) peak may reflect stimulus familiarity (34) and training (35); it is quite late for encoding bottom-up acoustic features but appropriately positioned to reflect a representation of auditory object formation (31). In this study, we investigate how amplitudes and latencies of each of these processing stages are affected by aging and listening condition in a continuous speech paradigm, to better understand the age-related temporal processing differences.

In comparison to the neural representations of the attended speech envelope discussed above, neural representations of unattended speech in older listeners are less well understood (although see Ref. 36). In this study, we extend the timing and fidelity analysis to unattended speech to understand how aging affects auditory scene representation in general, and stream segregation in particular, in the cortex.

In summary, this study aims to systematically investigate age-related neurophysiological effects on continuous speech processing, using both envelope reconstruction and TRF analysis. The effects of age, selective attention, and competing speech masking are evaluated concurrently. To minimize the effect of age-related peripheral hearing loss, only subjects who had hearing thresholds within clinically normal limits through 4.0 kHz were recruited in the study. Expanding beyond prior results, we expected to observe

exaggerated representation of the speech envelope in older adults irrespective of the competing speech, selective attention, and processing stage. As older adults show difficulty understanding speech in noise, we expect that timing and fidelity measures will be differently or severely affected from speech in quiet to speech in noise in older adults compared with younger adults. Finally, for the late neural processing stages, we expect that older adults' neural measures patterns will grow from those of younger adults (e.g., enhanced top-down connectivity associated with cognitive compensation in older adults).

METHODS

All experiments were performed in accordance with the guidelines and regulations for human subject testing of the University of Maryland's Institutional Review Board. All subjects gave written informed consent to participate in the protocol, and they were compensated for their time.

Subjects

Thirty-four native English speakers participated in the study: 18 younger adults (7 males; mean age 20 yr, range 18–26 yr) and 16 older adults (5 males; mean age 70 yr, range 65–78 yr). Data from two additional subjects (1 older and 1 younger) were not included in the analysis because of data saturation caused by excessive MEG artifacts. All subjects had normal hearing (see Fig. 1), defined as pure-tone thresholds ≤ 25 dB hearing level (HL) from 125 to 4,000 Hz in at least one ear and no more than 10 dB difference between ears at each frequency. Only subjects with Montreal Cognitive Assessments (MoCA) scores within normal limits (≥ 26) and no history of neurological disorder were included.

Stimuli and Experimental Design

The speech segments were extracted from the audio book, *The Legend of Sleepy Hollow*, by Washington Irving, narrated by separate male (<https://librivox.org/the-legend-of-sleepy-hollow-by-washington-irving>) and female (<https://www.amazon.com/The-Legend-of-Sleepy-Hollow-audiobook/dp/B00113CMHE>) talkers. Talker pauses > 400 ms were shortened to 400 ms, and then recordings were low-pass filtered < 4 kHz with a third-order elliptic filter. One-minute-long audio stimulus segments (22.05 kHz sampling rate) were extracted, and root mean square (rms) value of the sound amplitudes was normalized to have equal perceptual loudness in MATLAB (MathWorks) (32). All stimuli were presented diotically (identically in each ear).

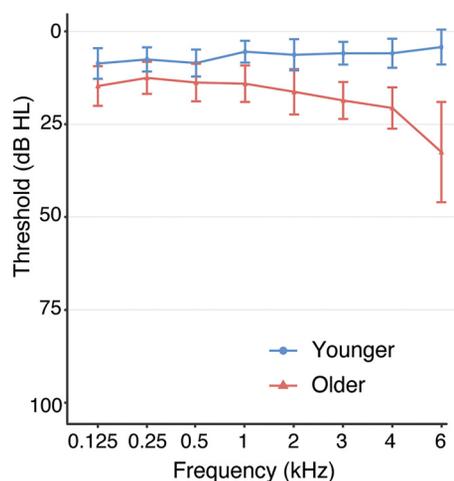


Figure 1. Audiogram of the grand average across ears for younger (blue) and older (red) subjects. Error bars indicate ± 1 SD. Both groups have clinically normal hearing, defined as pure-tone thresholds ≤ 25 dB hearing level (HL) from 125 to 4,000 Hz in at least 1 ear and no more than 10 dB difference between ears at each frequency.

amazon.com/The-Legend-of-Sleepy-Hollow-audiobook/dp/B00113CMHE) talkers. Talker pauses > 400 ms were shortened to 400 ms, and then recordings were low-pass filtered < 4 kHz with a third-order elliptic filter. One-minute-long audio stimulus segments (22.05 kHz sampling rate) were extracted, and root mean square (rms) value of the sound amplitudes was normalized to have equal perceptual loudness in MATLAB (MathWorks) (32). All stimuli were presented diotically (identically in each ear).

Four types of stimuli were presented: single talker (“quiet speech”), two talkers (“competing talkers”) at two different relative loudness levels [0 dB signal-to-noise ratio (SNR) and -6 dB SNR], and single talker mixed with three-talker babble (“babble speech”). For the competing-talker speech trials (a trial was defined as a 60-s-duration speech passage presentation), subjects were asked to selectively attend to one talker while ignoring the other, for which there were two SNR levels, 0 dB and -6 dB. For the babble condition, only the female talker was used as foreground, with the three-talker babble mixed in at 0 dB SNR. In the mixed speech and babble speech conditions, the sound level of the attended talker was identical to that of the corresponding single-talker condition; only the sound level for the unattended talker or babble was altered to change the noise condition. The order of the four competing-talker blocks was counterbalanced across subjects in the order of attended talker and SNR (2×2). The babble speech condition was always presented as the third block. In the competing-talker and babble speech conditions, each stimulus was presented three times. At the end of each of these blocks, the attended and unattended speech stimuli in that block were presented alone as single-talker speech without repetition (for the babble speech condition, only the attended female talker speech was presented); otherwise, no speech segment was ever reused across blocks.

Sound level was calibrated to ~ 70 dBA sound pressure level (SPL) with 500-Hz tones and equalized to be approximately flat from 40 Hz to 4 kHz. The stimuli were delivered with E-A-RTONE 3A tubes (impedance 50 Ω), which strongly attenuate frequencies above 3–4 kHz, and E-A-RLINK (Etymotic Research, Elk Grove Village, IL) disposable earbuds inserted into the ear canals.

To motivate the subjects to engage in the task, at the end of each trial a simple story-content question based on the attended passage was asked. After the first trial of each condition, subjects were also asked to rate the intelligibility rating on a scale from 0 to 10 (0 being completely unintelligible and 10 being completely intelligible). This estimated rating was used as a subjective measure of intelligibility.

Data Recording

Noninvasive neuromagnetic responses were recorded with a 157-axial gradiometer whole head MEG system (KIT, Kanazawa, Japan), inside a dimly lit, magnetically shielded room (Vacuumschmelze GmbH & Co. KG, Hanau, Germany) at the Maryland Neuroimaging Center. The data were sampled at 2 kHz along with an online antialiasing low-pass filter with cutoff frequency at 500 Hz and a 60-Hz notch filter. Three separate channels functioned as environmental

reference channels. Subjects lay supine during the entire experiment and were asked to minimize body movements. During the task, subjects were asked to keep their eyes open and fixate on a male/female cartoon face at the center of screen, corresponding to the attended talker. Pupil size data were recorded simultaneously with an eye tracker (EyeLink 1000 Plus); those results will be presented separately.

Data Processing

All data analysis was conducted in MATLAB R2020a. The raw MEG data were first denoised by removal of non-functioning and saturated channels, then with time-shifted principal component analysis (TSPCA) (37), using the three reference channels to project out environmental magnetic noise not related to the brain activity, and then with sensor noise suppression (SNS) (38) to project out sensor specific noise. To focus on low-frequency cortical activity, the remaining data were band-pass filtered between 1 and 10 Hz with an order 6000 Hamming-windowed finite impulse response filter (FIR) and compensated for group delay. A blind source separation method, denoising source separation (DSS) (39, 40), was next applied to the repeated trials to extract those subject-specific spatial components that are reliable over trials, ranked in order of reproducibility. The first six DSS components were used for the stimulus reconstruction analysis. Generally, the first component corresponds to the primary auditory component (i.e., with a topography consistent with bilateral temporal lobe sources) and so was selected for all subsequent TRF estimation; for one older adult, the second component reflected the primary auditory component and so was selected for TRF estimation in place of the first. As the polarity of a raw DSS component is arbitrary, which might affect TRF peak estimation and comparison between subjects, we aligned the polarity of the primary auditory component across subjects by flipping both the sign and the topography of that component for any subject whose topography gave a negative spatial correlation with a fiducial topography (i.e., the topography produced by an evoked response to a pure tone with standard M50/M100 peaks). Finally, data were downsampled to 250 Hz for TRF analysis and to 100 Hz for the stimulus reconstruction analysis.

The envelope of the audio waveform was processed to match the processed MEG data. Each attended and unattended single-talker stimulus was downsampled to 2 kHz, and the logarithmic envelope was extracted as described in Ref. 41. Then the envelope was filtered with the same band-pass filter (1–10 Hz) applied to the MEG data (and group delay compensation) and downsampled to 250 Hz and 100 Hz for TRF and stimulus reconstruction analysis, respectively.

Behavioral Tests

Flanker test.

The ability to selectively attend to one talker and ignore (inhibit) the other requires executive function. The Flanker Inhibitory Control and Attention Test of the National Institutes of Health Toolbox (42) was used as a measure of the subject's general behavioral ability to suppress competing stimuli in a visual scene. Subjects were instructed to identify

the direction of a central arrow while ignoring the directions of a surrounding set of four arrows (“flankers”) by pressing a key as quickly and accurately as possible. The direction of the central arrow could be similar to (congruent) or different from (incongruent) the surrounding arrows. The unadjusted scale score was calculated based on the reaction times (RTs) and the accuracy. Higher flanker scores represent better performance.

Speech-in-noise test.

A material-specific objective intelligibility test, referred to as the speech-in-noise (SPIN) task, was done on a separate day after the MEG recordings. Because of the COVID-19 pandemic, only data from 32 subjects were obtained: 18 younger adults and 14 older adults. The task was run via a graphical user interface in MATLAB. Subjects listened to 3- to 5-s-duration short sentence segments (with 4–7 key words) from the same audio book used for the MEG study, using different segments from those used in the MEG study (but processed identically). Subjects were asked to repeat back the speech segment in the case of quiet speech and the selectively attended speech segment otherwise. At each noise condition there were six different speech segments; the first segment was used as a practice trial and was not included in the accuracy calculation. The accuracy per each condition and talker was calculated as the ratio (number of correctly repeated key words)/(total number of key words per condition and talker). The same conditions (quiet speech, 0 dB SNR, –6 dB SNR, and babble speech; attend male and attend female) were used, presented in the same order as in the MEG study for that subject.

Quick Speech-in Noise test.

The Quick Speech-in-Noise (QuickSIN) test (43), a standardized measure of a listener's ability to understand speech in noise, was also employed. Because of the COVID-19 pandemic only data from half of the subjects were obtained; and therefore these data were not further analyzed.

Data Analysis

Stimulus (envelope) reconstruction.

Reconstruction of the speech envelope (backward/decoding model) from the neural response is a measure of cortical representation of the perceived speech. The low-frequency envelopes from the attended (foreground; att) and unattended (background; unatt) talkers are denoted by $E_{\text{att}}(t)$ and $E_{\text{unatt}}(t)$, respectively. For each subject and each trial, the attended and unattended speech envelopes were reconstructed separately (but simultaneously) with a linear temporal decoder applied to the first six DSS components [$D(d, t)$] estimated by the boosting algorithm (33, 44, 45) as follows:

$$E_{\text{acoustic}}(t) = \sum_{d=1}^6 \sum_{\tau=0}^T h(d, \tau) D(d, t + \tau) + \epsilon(t)$$

where $\epsilon(t)$ is the contribution not explained by the model and $h(d, \tau)$ is the decoder matrix value for component d at time lag τ . T is the integration window (500 ms unless specified otherwise). Tenfold cross-validation was used, resulting

in 10 decoding filters per trial. These 10 decoders were averaged to produce the final decoding filter for each trial. This filter was then used to reconstruct the speech envelope, and the decoder accuracy is given by the linear correlation coefficient between the reconstructed and the true speech envelope.

Integration window analysis.

Performing reconstruction analysis with a fixed 500-ms integration window does not provide any access to temporal processing details within that window. Employing different time intervals allows incorporation of different information, and age-related differences in temporal processing should manifest as different trajectories for how envelope reconstruction accuracy builds up over time. Thus, the integration window size was also systematically varied from 10 ms to 610 ms with a step size of 50 ms. Generalized additive mixed models (GAMMs) were used to analyze the resulting time (integration window duration) series data.

Temporal response function.

Envelope reconstruction is a robust measure of how well a neural response tracks the stimulus, but any such backward model must necessarily integrate over information regarding neural response time and sensors (46). In contrast, the TRF, which as a forward model relates how the neural responses were generated from the stimulus, allows interpretation of the stimulus-driven brain responses (32, 47), since it instead integrates over stimulus time, not response time.

TRF analysis, a linear method widely used to analyze the temporal processing of the auditory signal, predicts how the brain responds to acoustic features with respect to time. Additionally, a simultaneous two-talker TRF model uses the envelopes from both the foreground and background talkers, denoted by $E_{\text{att}}(t)$ and $E_{\text{unatt}}(t)$, respectively, with the model formulated as

$$r(t) = \sum_{\tau} h_{\text{att}}(\tau) E_{\text{att}}(t-\tau) + \sum_{\tau} h_{\text{unatt}}(\tau) E_{\text{unatt}}(t-\tau) + \varepsilon(t)$$

where $r(t)$ is the cortical response at a particular sensor, τ is the time lag relative to the speech envelope $E(t)$, and $\varepsilon(t)$ is the residual cortical response not explained by the linear model. $h_{\text{att}}(\tau)$ and $h_{\text{unatt}}(\tau)$ describe the filters that define the linear neural encoding from speech envelope to the neural response and are known as the TRFs for the attended and unattended speech, respectively. The range of τ is chosen to range from 0 to 500 ms. The competing TRFs were estimated simultaneously with the boosting algorithm with 10-fold cross-validation, to minimize the mean square error between the predicted neural response and the true neural response (44, 48). For the babble condition, the summed three-talker babble speech envelope was used as the background. The final TRF was evaluated as the mean TRF over the 10-fold cross-validation sets. TRFs were estimated for each condition and subject on the concatenated data, giving one TRF per subject and condition (e.g., in the 0 dB case, all 6 such trials were concatenated before applying the boosting algorithm).

Larger amplitudes in the TRF indicate that the neural populations with the corresponding latencies follow the speech envelope better when synchronously responding to the

stimulus. The TRF has three prominent peaks, with latencies at ~ 50 ms (positive peak), ~ 100 ms (negative peak), and 200 ms (positive peak), named the $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$, respectively. Each peak corresponds to a different stage in the auditory signal processing chain. Latency and polarity of these peaks under different conditions can be compared similarly to the P1, N1, and P2 CAEP peaks. For example, an age- or task-related increase in $M50_{\text{TRF}}$ amplitude represents a stronger response at that latency under identical stimulus conditions. For each subject and condition, peak latencies were extracted within a specific time range; for $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$ the windows were 30–110 ms, 80–200 ms, and 140–300 ms, respectively. These peak amplitudes and latencies were further analyzed to evaluate the effects of aging, task difficulty, and selective attention.

Statistical Analysis

All statistical analysis was performed in R version 4.0 (49). Linear mixed-effect models (LMMs) were used to systematically evaluate the relationships between the dependent (behavioral scores, neural measures) and independent (age, noise condition, selective attention) variables. For the LMM analysis, the *lme4* (version 1.1-30) (50) and *lmerTest* (version 3.1-30) (51) packages in R were used. The best-fit model from the initial full model was found by the *buildmer* package (version 2.4) (52) using the default settings, where the *buildmer* function first determines the order of parameters based on the likelihood-ratio test (LRT) and then uses a backward elimination stepwise procedure to identify the model of best fit for random and fixed effects. The assumptions of mixed-effect modeling, linearity, homogeneity of variance, and normality of residuals, were checked per each best-fit model based on the residual plots. Reported β values represent changes in the dependent measure when comparing one level of an independent variable versus its reference level. P values were calculated with the Satterthwaite approximation for degrees of freedom (53, 54). To interpret significant fixed-effect interaction terms, variables were relevelled to obtain model estimates for individual factor levels (indicated by “with [new reference level] reference level”). The summary tables for each model used in RESULTS are reported in the Supplemental Materials (available at <https://doi.org/10.13016/4goq-vs0z>).

The subjective intelligibility ratings and objective intelligibility scores (SPIN scores) were analyzed separately, each using a LMM with age as a between-subject factor (categorical variable; 2 levels: younger [reference level], older) and noise condition as a within-subject factor (categorical variable; 4 levels: quiet [reference level], 0 dB, –6 dB, and babble). To account for repeated measures, we used subject as a clustering variable so that the intercept and effects of noise condition could vary across subjects (random intercept for subject and random slopes for noise condition by subject respectively). The full model for each dependent variable was defined as intelligibility \sim age \times noise condition + (noise condition|subject). To evaluate the relationship between the two measures, data were analyzed with a separate LMM with SPIN score as the dependent variable, intelligibility rating as the independent variable, and subject as a random intercept: SPIN score \sim intelligibility rating + (1|subject).

LMMs were used to systematically evaluate the relationships among the computed neural features (reconstruction accuracy; $M50_{TRF}$, $M100_{TRF}$, $M200_{TRF}$ for both amplitude and latency) and age, noise condition, and selective attention (categorical variable; 2 levels: attended [reference level], unattended). Two models were generated for each neural feature 1) to examine the effects of aging and noise condition on the neural features measured for the attended talker and 2) to examine the effects of aging, noise condition, and attention on the neural features measured for the attended talker and the unattended talker. Only data from two competing-talker noise conditions were used for the latter model. The full models for 1 and 2 were defined as neural feature \sim age \times noise condition + (1 + noise condition|subject) and neural feature \sim age \times noise condition \times attention + (1 + noise condition \times attention|subject), respectively.

To model nonlinear changes in the integration window analysis, generalized additive mixed models (GAMMs) (55) in R [packages *mgcv* (version 1.8-40), *itsadug* (version 2.4.1)] were used to analyze the reconstruction accuracies over integration window. Compared with generalized linear mixed-effect models, GAMMs have several advantages for modeling time series data, especially in electrophysiology (56, 57) and pupillometry (58). In particular, they 1) can model both linear and nonlinear patterns in the data using parametric and smooth terms and 2) can, critically, include various types of autoregressive (AR) correlation structures that deal with autocorrelational structure in the errors. Compared with conventional nonlinear mixed-effect modeling, where nonlinear trends are fitted by polynomials of the predictor, GAMMs fit with p-spline based “smooth terms” with a specified number of basis functions (knots) that specify how “wiggly” the model can be. For each fixed-effect term (age, noise condition, and attention) and corresponding smooth term, model testing was compared between a test model and baseline model with χ^2 (*compareML* in the *itsadug* package) to determine the significance of predictors (58, 59). The parametric term indicates the overall difference in height between two curves, whereas smooth terms indicate the difference in shape (i.e., “wiggleness”) between two curves. The models included random smooths for each subject to capture the individual trends in reconstruction accuracies over integration window. For the time series reconstruction accuracies here, since the current time point was observed to be correlated with the next time point, we employed an autoregression model 1 (AR1) structure. The assumptions of mixed-effect modeling and model diagnostics (oversmooth or undersmooth) were performed by residual plots and the *gam.check()* function.

Code and Data Availability

Preprocessed MEG data, stimulus material, behavioral data, and analysis codes are available at <https://doi.org/10.13016/4goq-vs0z>.

RESULTS

Behavioral Data

Flanker test.

The effect of age on the flanker scores was analyzed with a two-sample *t* test. Results showed significantly better scores

in younger adults than in older adults ($t_{32} = 6.0956$, $P < 0.001$), suggesting that older adults' performance in inhibition task may decline with aging.

SPIN test.

The effects of age and noise condition on the objective intelligibility measures (SPIN scores) were analyzed with LMM. Figure 2A plots the SPIN scores for both groups at all noise conditions. The best-fit model included main effects of noise condition on speech SPIN scores and random intercept by subject (Supplemental Table S1). There was no effect of age on the SPIN scores. SPIN scores significantly dropped from quiet speech to every other condition (Quiet to 0 dB to -6 dB to Babble), with the highest drop from 0 dB to -6 dB [with Quiet reference level: noise condition(0 dB): $\beta = -12.88$, SE = 1.57, $P < 0.001$; with 0 dB reference level: noise condition(-6 dB): $\beta = -40.95$, SE = 1.57, $P < 0.001$; with -6 dB reference level: noise condition(Babble): $\beta = -13.52$, SE = 1.92, $P < 0.001$]. The lack of significant age effects is addressed in DISCUSSION.

Intelligibility ratings.

Parallel LMM analysis of the subjective intelligibility ratings (Fig. 2B) revealed fixed effects of both age and noise condition along with random slopes for noise condition by subject (Supplemental Table S2). Older adults rated the intelligibility slightly higher compared with younger adults ($\beta = 0.76$, SE = 0.36, $t = 2.11$, $P = 0.04$). As in the case of SPIN scores, intelligibility ratings dropped significantly from quiet speech to 0 dB SNR to -6 dB to Babble noise condition in both groups [with Quiet reference level: noise condition(0 dB): $\beta = -2.6$, SE = 0.27, $P < 0.001$; with 0 dB reference level: noise condition(-6 dB): $\beta = -1.51$, SE = 0.27, $P < 0.001$; with -6 dB reference level: noise condition(Babble): $\beta = -0.53$, SE = 0.34, $P = 0.13$].

To examine the consistency of the two measures, a separate LMM model was constructed to predict SPIN scores from intelligibility ratings. The best-fit model revealed that the subjective intelligibility ratings were positively related to the objective intelligibility scores ($\beta = 8.2$, SE = 0.58, $P < 0.001$) and revealed no effects of age.

Stimulus (Envelope) Reconstruction Analysis

As a simpler precursor to the full TRF analysis, we employed reconstruction analysis to investigate how the cortical representation of the speech envelope is affected by aging at a coarser level. First, we investigated the effects of age and noise condition on the attended talker envelope tracking. As summarized in Supplemental Table S3, the best-fit model revealed main effects of age and noise condition and age \times noise condition interactions with random intercepts by subject. For both groups and for all noise conditions, the reconstruction accuracies fitted by the model for the attended talker are shown in Fig. 3A. The main effects of age revealed that aging is associated with higher reconstruction accuracies ($\beta = 0.04$, SE = 0.01, $P < 0.001$) in all noise conditions.

The significant interaction age \times noise condition term revealed that aging adversely affects speech reconstruction from quiet to noisy speech [age(Older) \times noise condition(0 dB): $\beta = -0.018$, SE = 0.01, $P = 0.01$; age(Older) \times noise condition(-6 dB): $\beta = -0.015$, SE = 0.01, $P = 0.03$; age(Older) \times

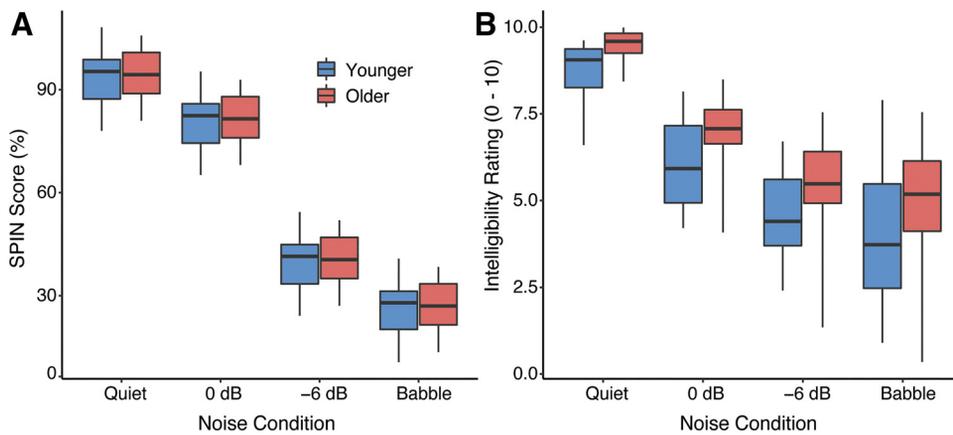


Figure 2. Model-predicted behavioral test results for speech-in-noise (SPIN) scores (0–100%) (A) and intelligibility ratings (0–10) (B). Both scores drop significantly with the noise condition. A significant effect of age was seen only for the intelligibility rating, whereas no age effect was found on the SPIN scores.

noise condition(Babble): $\beta = -0.023$, SE = 0.01, $P = 0.01$]. However, this effect was not significant from noisy speech to babble speech [with 0 dB reference level: age(Older) \times noise condition(Babble): $\beta = -0.005$, SE = 0.01, $P = 0.58$]. As can be seen from Fig. 3A, the main effect of noise condition revealed that the attended talker envelope reconstruction accuracies significantly reduce from quiet to noisy conditions in both groups [noise condition(0 dB): $\beta = -0.02$, SE = 0.004, $P < 0.001$; noise condition(-6 dB): $\beta = -0.02$, SE = 0.004, $P < 0.001$; noise condition(Babble): $\beta = -0.05$, SE = 0.005, $P < 0.001$]. No significant difference was observed between the 0 dB and -6 dB noise conditions [with 0 dB reference level: noise condition(-6 dB): $\beta = 0.003$, SE = 0.004, $P = 0.39$]. However, reconstruction accuracies significantly dropped from noisy speech to babble speech [with 0 dB reference level: noise condition(Babble): $\beta = -0.03$, SE = 0.005, $P < 0.001$; with -6 dB reference level: noise condition(Babble): $\beta = -0.03$, SE = 0.005, $P < 0.001$].

Second, to investigate the combined effects of selective attention, age, and noise condition, a separate analysis was performed on only the competing-talker speech data (0 dB and -6 dB) by including both attended and unattended speech envelope reconstruction accuracies. As shown in

Fig. 3B, LMM analysis revealed main effects of age and selective attention on reconstruction accuracy, with both random intercepts and random slopes for selective attention, by subject (Supplemental Table S4). Results revealed that the cortical representation of the speech envelope as measured by reconstruction accuracy is enlarged/overrepresented in older adults for both attended and unattended talkers ($\beta = 0.03$, SE = 0.01, $P < 0.001$). Furthermore, in both groups the attended talker was better represented than the unattended talker ($\beta = -0.03$, SE = 0.004, $P < 0.001$). For attended talker versus unattended talker, or for either of the age groups, no significant difference was observed between the 0 dB and -6 dB noise conditions.

Integration Window Analysis

Integration window analysis was done with GAMMs including both quiet and two-talker mixed speech for both attended and unattended speech envelope reconstruction accuracies. The initial model included a smooth term over the integration window, characterizing the nonlinearity of these functions. Model comparisons determined that separate smooths for age \times noise condition \times attention significantly improve the model fit [$\chi^2(23) = 212.1$, $P < 0.001$]. Moreover, adding the fixed effect term, characterizing the

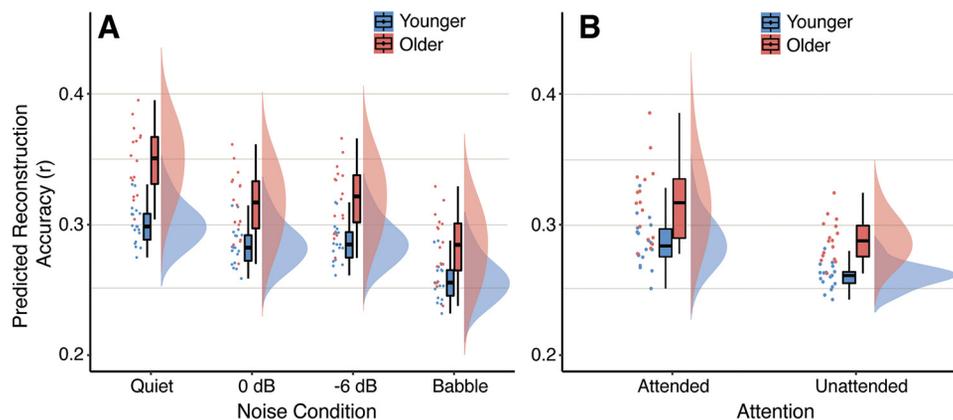


Figure 3. Model-predicted values for reconstruction accuracy for the attended speech for both younger and older adults and for all noise conditions (A) and the attended vs. unattended speech envelope reconstruction accuracies for competing talker conditions (B) (there is no separation by noise condition since no significant dependence on noise condition was found). Both attended and unattended speech envelope reconstructions illustrate that the speech reconstruction was significantly higher in older adults. When a single or multiple competing talkers are added, attended talker reconstruction accuracies significantly decrease in both groups. In both groups attended talker envelope reconstruction was higher compared with unattended.

height of these functions, age \times noise condition \times attention, and random smooth per subjects, also improved the model fit [$\chi^2(32) = 1,531.39$, $P < 0.001$ and $\chi^2(34) = 1,886.42$, $P < 0.001$, respectively]. Finally, as the residual plots showed a high autocorrelation in the residual analysis, an autoregression (AR1) model was included. The statistical information on this model (parametric terms and smooth terms) is summarized in Supplemental Table S10. **Figure 4A** shows the resulting smooth plots for the two groups. The results show that envelope reconstruction accuracy initially rapidly increases as the time window duration increases and then stabilizes to a slower rate as longer latencies are included.

To investigate how the integration window affects selective attention effects, the difference between attended and unattended talker responses were analyzed. **Figure 4B** shows the dynamical differences between attended and unattended talker reconstruction accuracy curves for both younger and older adults and for the 0 dB noise condition. The color-coded horizontal lines at the bottom of the graph indicate where the differences are significant. In both age groups, attended talker reconstruction accuracies were significantly higher compared with unattended talker after the middle processing stage (~ 150 ms). Interestingly, in older adults, for the -6 dB noise condition the unattended talker representation is enhanced compared with the attended talker during the early processing stages (~ 50 ms). Difference analysis emphasizing age group differences revealed that older adults' overrepresentation of the speech envelope starts as early as ~ 85 ms for the attended (**Fig. 4C**) and 55 ms for the unattended talker when averaged across noise conditions. As can be seen from **Fig. 4C**, the difference monotonically increases until ~ 300 ms and then levels off, suggesting that older adults' rate of increase in reconstruction accuracy with integration window is higher compared with younger adults until later processing stages.

Temporal Response Function Analysis

Unlike stimulus reconstruction analysis, which integrates over latencies, TRF analysis allows direct analysis of neural processing associated with any latency. For each TRF component ($M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$) we separately analyzed individuals' peak amplitude and latency, comparing between the two age groups and for both attended and unattended talker and noise conditions. **Figure 5A** visualizes how the TRF peak amplitudes and latencies varied for the attended talker across two age groups and noise conditions. **Figure 6A** visualizes how the TRF peak amplitudes and latencies varied between attended and unattended talker.

TRF peak amplitudes.

LMMs were fitted to the $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$ amplitudes separately to analyze the effect of age and noise condition on the attended talker TRF peak amplitudes for each neural processing stage. The best-fit model indicated main effects of age and noise condition, and an age \times noise condition interaction along with random intercept by subject for all three peaks $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$ (Supplemental Table S5).

TRF peak amplitudes for the $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$ are plotted in **Fig. 5, A** and **B**. Overall, in the comparison between younger and older, older adults showed exaggerated peak amplitudes in all three processing stages [$M50_{\text{TRF}}$: age(Older): $\beta = -0.01$, SE = 0.004, $P = 0.03$; $M100_{\text{TRF}}$: age(Older): $\beta = -0.02$, SE = 0.005, $P = 0.008$; and $M200_{\text{TRF}}$: age(Older): $\beta = -0.018$, SE = 0.003, $P = 0.001$]. For both the $M50_{\text{TRF}}$ and $M200_{\text{TRF}}$, peak amplitudes were stronger in all noise conditions, and that was significant only for the quiet speech condition [$M50_{\text{TRF}}$: age(Older): $\beta = 0.02$, SE = 0.005, $P < 0.001$; $M200_{\text{TRF}}$: age(Older): $\beta = 0.03$, SE = 0.01, $P < 0.001$]. In contradistinction, except for the quiet speech, the $M100_{\text{TRF}}$ was stronger in all the noisy conditions [age(Older):

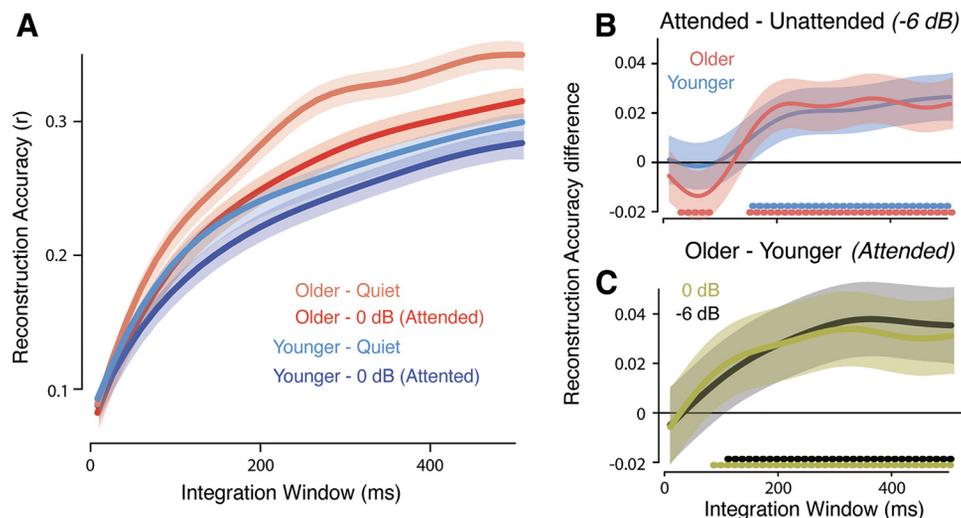


Figure 4. Integration window analysis using generalized additive mixed models (GAMMs). **A:** changing reconstruction accuracy with integration window (only a subset of curves is shown, for visual clarity). **B:** reconstruction accuracy difference between attended and unattended talker for -6 dB noise condition. **C:** reconstruction accuracy difference between older and younger for attended talker. The color-coded horizontal lines above the horizontal axis in **B** and **C** mark where the difference is statistically significant. The shaded area represents the 95% confidence interval (CI). In both groups reconstruction accuracy initially increases rapidly with increasing integration window, slowing down after ~ 300 ms. The attended talker is better represented than the unattended after ~ 170 ms. The overrepresentation of the attended talker envelope starts at early processing stages (< 100 ms).

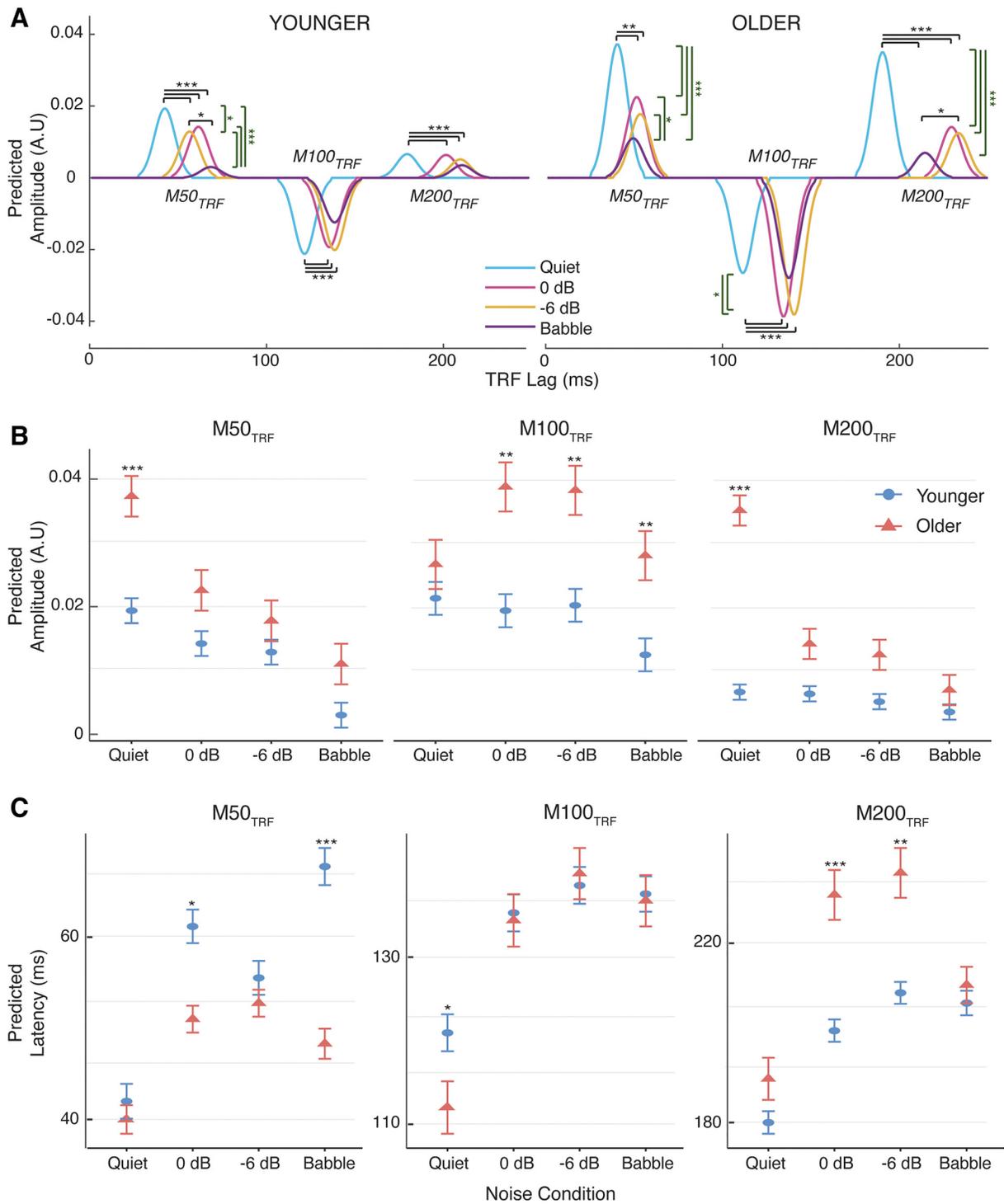


Figure 5. Model-predicted attended talker temporal response function (TRF) peak amplitudes and latencies. **A:** TRFs showing overall amplitudes and latency for both groups and all noise conditions (for visualization simplicity, all peaks are represented with the same Gaussian shape, SD 7 ms, centered at the group mean peak latency, and with amplitude given by the group mean peak amplitude). M50_{TRF}, M100_{TRF}, M200_{TRF}, peaks with latencies of ~50, ~100, and ~200 ms, respectively. **B:** the TRF amplitudes (\pm SE) as a function of noise condition (for the M100_{TRF}, as a negative polarity peak, the unsigned magnitude is shown). Generally, older adults (red) exhibit stronger TRF peak amplitudes compared with younger adults (blue). When a competing talker is added to the stimulus, the attended M50_{TRF} amplitudes decrease in both groups. From the quiet to competing speech conditions, M100_{TRF} increases and M200_{TRF} decreases but only in older adults. **C:** the TRF latencies (\pm SE) as a function of noise condition. Compared with younger adults, in older adults the M50_{TRF} is earlier and the M200_{TRF} is later. With task difficulty, peaks are typically delayed in both groups, with some exceptions in the babble condition. A.U., arbitrary units. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

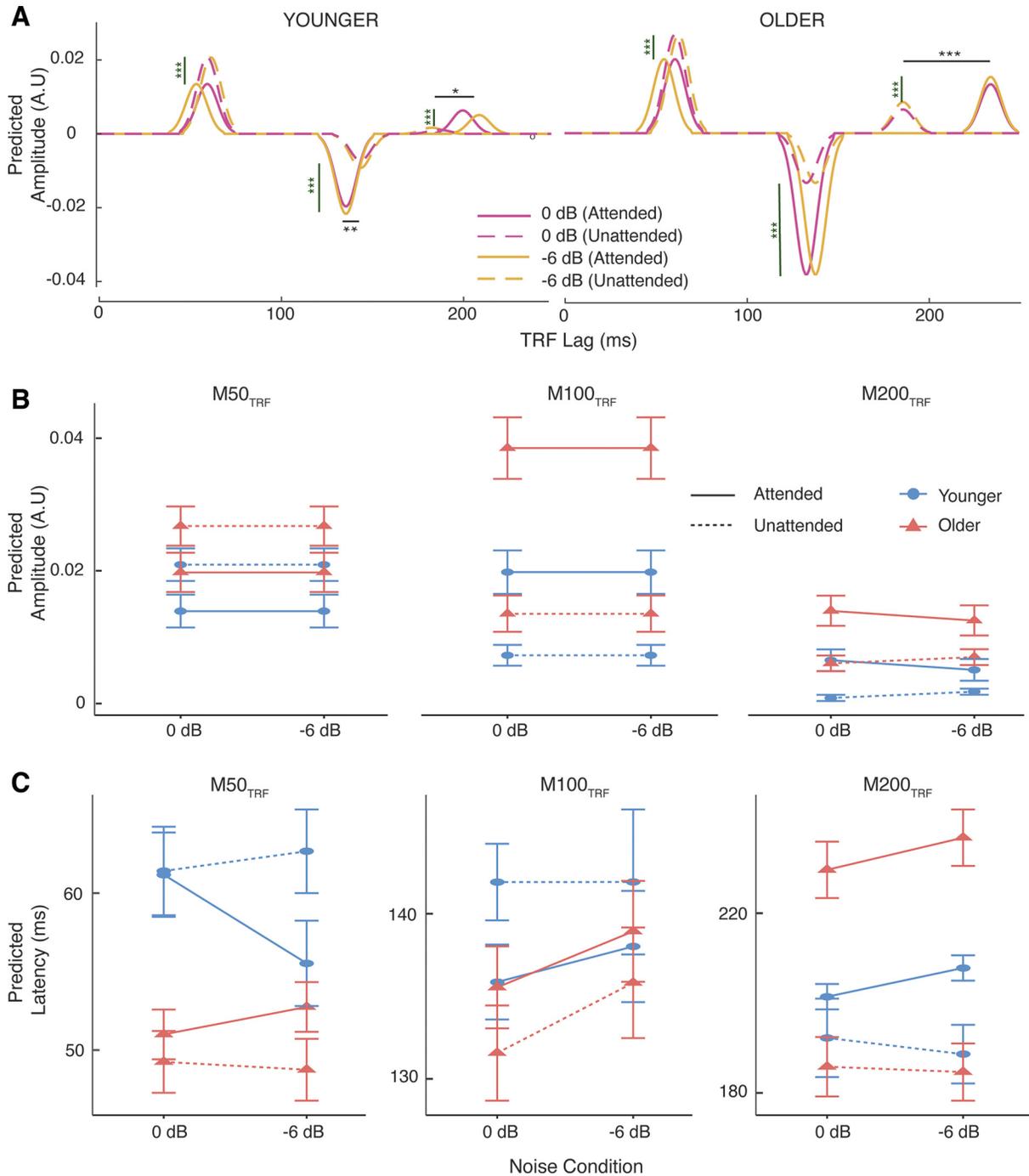


Figure 6. Model-predicted values for attended vs. unattended talker temporal response function (TRF) peak amplitude and latency. **A:** TRF peak amplitudes and latency for both groups and two-talker speech conditions for both attended and unattended talker (solid line, attended; dashed line, unattended). For visualization simplification, peaks are represented with a common Gaussian shape as in Fig. 5. **B:** TRF amplitudes (\pm SE) as a function of noise condition (for the M100_{TRF}, as a negative polarity peak, the unsigned magnitude is shown). M50_{TRF}, M100_{TRF}, M200_{TRF}, peaks with latencies of \sim 50, \sim 100, and \sim 200 ms, respectively. Generally, older adults exhibit stronger TRF amplitudes for both attended and unattended peaks. Attended M50_{TRF} is significantly smaller compared with unattended amplitudes. In contrast, attended M100_{TRF} and M200_{TRF} are enhanced compared with unattended peak amplitude. Note that the model analysis for the M50_{TRF} and M100_{TRF} amplitude did not find significant differences between noise conditions, so the model mean and SE do not change there. **C:** TRF latencies (\pm SE) as a function of noise condition. The attended M200_{TRF} peak is significantly delayed compared with the unattended M200_{TRF} peak, and this difference is bigger in older adults. A.U., arbitrary units. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

$\beta = 0.01$, SE = 0.01, $P = 0.036$; with 0 dB reference level: age (Older): $\beta = 0.02$, SE = 0.01, $P = 0.001$; with -6 dB reference level: age(Older): $\beta = 0.02$, SE = 0.01, $P = 0.002$; with Babble reference level: age(Older): $\beta = 0.02$, SE = 0.01, $P = 0.008$].

The effects of noise condition revealed that the M50_{TRF} response decreases from quiet to every other noise condition in both groups [noise condition(0 dB): $\beta = -0.01$, SE = 0.003, $P = 0.01$; noise condition(-6 dB): $\beta = -0.01$, SE = 0.003, $P =$

0.04; noise condition(Babble): $\beta = -0.02$, $SE = 0.003$, $P < 0.001$; with Older reference level: noise condition(0 dB): $\beta = -0.01$, $SE = 0.003$, $P < 0.001$; noise condition(-6 dB): $\beta = -0.2$, $SE = 0.003$, $P < 0.001$; noise condition(Babble): $\beta = -0.03$, $SE = 0.003$, $P < 0.001$]. A significant age \times noise condition interaction indicated that aging contributes more to the $M50_{TRF}$ reduction as a function of noise condition [age(Older) \times noise condition(0 dB): $\beta = -0.01$, $SE = 0.004$, $P = 0.03$]. In contrast to the $M50_{TRF}$, the $M100_{TRF}$ and $M200_{TRF}$ did not significantly vary across the quiet and two-talker noisy conditions in younger adults. However, in older adults, from quiet to two-talker noise conditions the $M100_{TRF}$ significantly increased [with Older reference level: noise condition(0 dB): $\beta = 0.01$, $SE = 0.003$, $P < 0.001$], whereas the $M200_{TRF}$ decreased [with Older reference level: noise condition(0 dB): $\beta = -0.02$, $SE = 0.003$, $P < 0.001$]. Interestingly, in both groups the $M100_{TRF}$ peak amplitudes significantly dropped from -6 dB to the babble condition [with -6 dB reference level: noise condition(Babble): $\beta = -0.01$, $SE = 0.003$, $P = 0.02$; with Older, -6 dB reference level: noise condition(Babble): $\beta = -0.01$, $SE = 0.003$, $P = 0.003$].

To investigate the combined effects of aging, selective attention, and noise condition on the TRF amplitude responses, separate LMM models were constructed (Supplemental Table S6). Figure 6B displays the TRF amplitude variation for the attended and unattended talker, for both age groups and for competing-talker conditions (0 dB and -6 dB). LMM applied to the $M50_{TRF}$ showed a main effect of attention, revealing that the unattended $M50_{TRF}$ amplitude is bigger compared with the attended $M50_{TRF}$ amplitude in both groups [attention(Unattended): $\beta = 0.01$, $SE = 0.001$, $P < 0.001$]. In contrast, the $M100_{TRF}$ peak amplitudes showed main effects of age and attention, and an age \times attention interaction effect. Compared with younger adults, both the attended and unattended $M100_{TRF}$ peak amplitudes were exaggerated in older adults; however, this effect was statistically significant only for the attended talker peak amplitudes [age(Older): $\beta = 0.02$, $SE = 0.01$, $P < 0.001$; with Unattended reference level: attention(Unattended): age(Older): $\beta = 0.01$, $SE = 0.01$, $P = 0.19$]. In both age groups unattended peak amplitudes were reduced compared with attended peak amplitudes [attention(Unattended): $\beta = -0.01$, $SE = 0.003$, $P < 0.001$; with Older reference level: attention(Unattended): $\beta = -0.02$, $SE = 0.003$, $P < 0.001$], and the interaction effect revealed that this reduction in peak heights is amplified by aging ($\beta = -0.01$, $SE = 0.004$, $P < 0.001$). Interestingly, the $M200_{TRF}$ showed main effects of age and attention, and an attention \times noise condition interaction. $M200_{TRF}$ amplitudes in both attended and unattended TRFs were stronger in older adults [age(Older): $\beta = 0.01$, $SE = 0.001$, $P < 0.001$]. The selective attention effect revealed that in both groups the attended talker $M200_{TRF}$ peak amplitude is stronger compared with the unattended [attention(Unattended): $\beta = -0.01$, $SE = 0.002$, $P < 0.001$].

TRF peak latencies.

Similar to TRF peak amplitude analysis, TRF peak latency analysis was performed on the attended talker TRFs as the first step. The best-fit model revealed effects of age, noise condition, and age \times noise condition along with random intercepts by subject for $M50_{TRF}$, $M100_{TRF}$, and $M200_{TRF}$ (Supplemental Table S7). Model-predicted latencies are

plotted in Fig. 5C. Averaged latencies over noise conditions revealed that compared with younger adults the older adults' early peak, $M50_{TRF}$, is significantly earlier [age(Older): $\beta = 8.4$, $SE = 3.1$, $P = 0.01$] and that there is no significant latency difference for the middle peak, $M100_{TRF}$ [age(Older): $\beta = 2.46$, $SE = 4.15$, $P = 0.55$] whereas the late peak $M200_{TRF}$ is significantly delayed [age(Older): $\beta = -17.2$, $SE = 6.37$, $P = 0.01$]. These results suggest that the three distinct processing stages are each differently affected by aging.

All three peaks were significantly delayed for the noisy conditions, relative to quiet, for both younger [$M50_{TRF}$: noise condition(0 dB): $\beta = 18.48$, $SE = 3.02$, $P < 0.001$; $M100_{TRF}$: noise condition(0 dB): $\beta = 14.59$, $SE = 2.75$, $P < 0.001$; $M200_{TRF}$: noise condition(0 dB): $\beta = 21.80$, $SE = 4.64$, $P < 0.001$] and older [with Older reference level: $M50_{TRF}$: noise condition(0 dB): $\beta = 10.83$, $SE = 3.09$, $P < 0.001$; $M100_{TRF}$: noise condition(0 dB): $\beta = 22.13$, $SE = 2.92$, $P < 0.001$; $M200_{TRF}$: noise condition(0 dB): $\beta = 38.80$, $SE = 4.86$, $P < 0.001$] adults, highlighting that the peak responses are delayed with the stimulus noise condition. With respect to quiet, babble speech latencies in all three processing were delayed in both younger [noise condition(Babble): $M50_{TRF}$: $\beta = 25.62$, $SE = 3.12$, $P < 0.001$; $M100_{TRF}$: $\beta = 16.99$, $SE = 2.71$, $P < 0.001$; $M200_{TRF}$: $\beta = 30.89$, $SE = 5.17$, $P < 0.001$] and older adults [with Older reference level: noise condition(Babble): $M50_{TRF}$: $\beta = 9.24$, $SE = 3.23$, $P = 0.006$; $M100_{TRF}$: $\beta = 25.63$, $SE = 2.87$, $P < 0.001$; $M200_{TRF}$: $\beta = 24.59$, $SE = 4.98$, $P < 0.001$] adults. The age \times noise condition interaction manifests as aging contributing more to the peak delay from quiet to 0 dB for both $M100_{TRF}$ [age(Older) \times noise condition(0 dB): $\beta = 8.05$, $SE = 3.98$, $P = 0.04$] and $M200_{TRF}$ [age(Older) \times noise condition(0 dB): $\beta = 17.08$, $SE = 6.68$, $P = 0.01$].

The effect of selective attention on TRF peak latencies was analyzed with LMMs (Supplemental Table S8), and results are displayed in Fig. 6C. Comparing mean latencies across noise conditions revealed that compared with younger adults unattended peaks are earlier in older adults for both the $M50_{TRF}$ and $M100_{TRF}$ [with Unattended reference level: age(Older): $M50_{TRF}$: $\beta = -12.96$, $SE = 3.96$, $P < 0.001$; $M100_{TRF}$: $\beta = -13.98$, $SE = 4.23$, $P = 0.003$]. This effect was, however, not significant for the $M200_{TRF}$ [with Unattended reference level: age(Older): $\beta = -0.70$, $SE = 8.65$, $P = 0.94$]. Differences due to selective attention on the neural response latencies were analyzed with the same LMM models. Results indicated that there is no significant latency difference between attended and unattended peaks for the early peak $M50_{TRF}$ in both groups [attention(Unattended): $\beta = 0.50$, $SE = 3.11$, $P = 0.87$; with Older reference level: attention(Unattended): $\beta = 2.88$, $SE = 2.83$, $P = 0.32$], whereas for the middle peak $M100_{TRF}$ the attended peak was earlier compared with unattended only in younger adults [attention(Unattended): $\beta = 8.03$, $SE = 3.21$, $P = 0.01$]. For the late peak $M200_{TRF}$, both age groups showed a delayed attended peak compared with the unattended peak [attention(Unattended): $\beta = -12.02$, $SE = 6.28$, $P = 0.05$], and this effect was stronger in older adults [age(Older) \times attention(Unattended): $\beta = -28.54$, $SE = 8.26$, $P = 0.001$].

Amplitude vs. latency analysis.

Potential associations between TRF peak amplitudes and latencies for the attended talker were analyzed for $M50_{TRF}$,

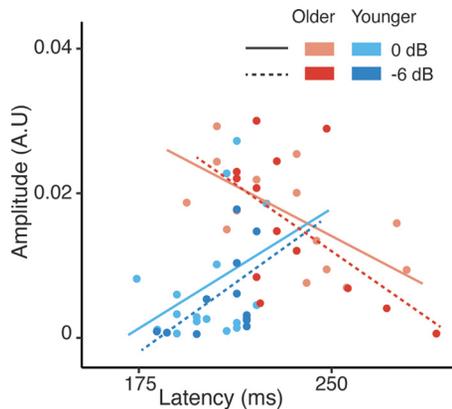


Figure 7. Temporal response function peak with latency of ~ 200 ms ($M200_{TRF}$) peak amplitude vs. latency for 0 dB and -6 dB conditions. Older adults' amplitudes were significantly negatively associated with their latencies; younger adults showed no significant association. A.U., arbitrary units.

$M100_{TRF}$, and $M200_{TRF}$ separately. The analysis was conducted on 0 dB and -6 dB noise conditions (quiet and babble conditions were excluded, as they represent two extreme acoustic environments resulting in heavily restricted dynamic range in both amplitude and latency). Peak amplitudes were predicted by age and peak latencies. As can be seen from Fig. 7, older adults' $M200_{TRF}$ peak amplitudes exhibited a significant negative relationship with latency [with Older reference level: latency: 0 dB condition: $\beta = -0.0001$, $SE = 0.0001$, $P = 0.02$; -6 dB condition: $\beta = -0.0002$, $SE = 0.0001$, $P = 0.01$], i.e., delayed peaks showed smaller peak amplitudes, but this was not seen for the earlier peaks. Conversely, in younger adults peak amplitudes were not related to the latencies (latency: 0 dB condition: $\beta = 0.0002$, $SE = 0.0001$, $P = 0.1$; -6 dB condition: $\beta = 0.0003$, $SE = 0.0001$, $P = 0.1$) (Supplemental Table S9).

Relationships among Neural Features and Behavioral Responses

LMMs were used to evaluate the relationship between the neural measures (reconstruction accuracies, TRF peak

amplitudes and latencies) and the behavioral measures. First, we analyzed how attended talker neural features are affected by speech intelligibility score and age but without specific regard to stimulus noise condition (noise condition and behavioral intelligibility measures are too correlated to include both measures; see Fig. 2A). Results revealed that the reconstruction accuracy increases with better speech intelligibility in both groups (SPIN score: $\beta = 0.0005$, $SE = 0.001$, $P < 0.001$) (Fig. 8A). Similarly, TRF peak amplitude analysis revealed that stronger $M50_{TRF}$ amplitudes are associated with better speech intelligibility score (SPIN score: $\beta = 0.0002$, $SE = 0.001$, $P < 0.001$) (Fig. 8B), and this effect was stronger in older adults [age(Older) \times SPIN score: $\beta = 0.0002$, $SE = 0.001$, $P = 0.01$]. However, no significant trends were found for $M100_{TRF}$ amplitude. Interestingly, for the late peak $M200_{TRF}$ older adults showed smaller peak amplitudes with poorer speech intelligibility score (with Older reference level: SPIN score: $\beta = 0.0003$, $SE = 0.001$, $P < 0.001$), whereas no significant trend was found for younger adults (SPIN score: $\beta = 0.0001$, $SE = 0.001$, $P = 0.022$) (Fig. 8C). Analysis of peak latencies revealed that, in both groups, peak latencies at all three stages are negatively related to the speech intelligibility score (SPIN score: $M50_{TRF}$: $\beta = -0.17$, $SE = 0.04$, $P < 0.001$; $M100_{TRF}$: $\beta = -0.22$, $SE = 0.03$, $P < 0.001$; $M200_{TRF}$: $\beta = -0.32$, $SE = 0.06$, $P < 0.001$). Additionally, similar trends were observed when subjective intelligibility rating was used instead of the SPIN scores.

However, when noise condition was added into the model, for any one noise condition no consistent trends were found between behavioral scores and neural measures.

DISCUSSION

This study examined the effects of aging on neural measures of cortical continuous speech processing under difficult listening conditions. These neural measures include envelope reconstruction, integration window analysis, and TRF analysis. The results were aligned with previous findings that aging is associated with exaggerated cortical representations of speech (10, 14) and further investigated the

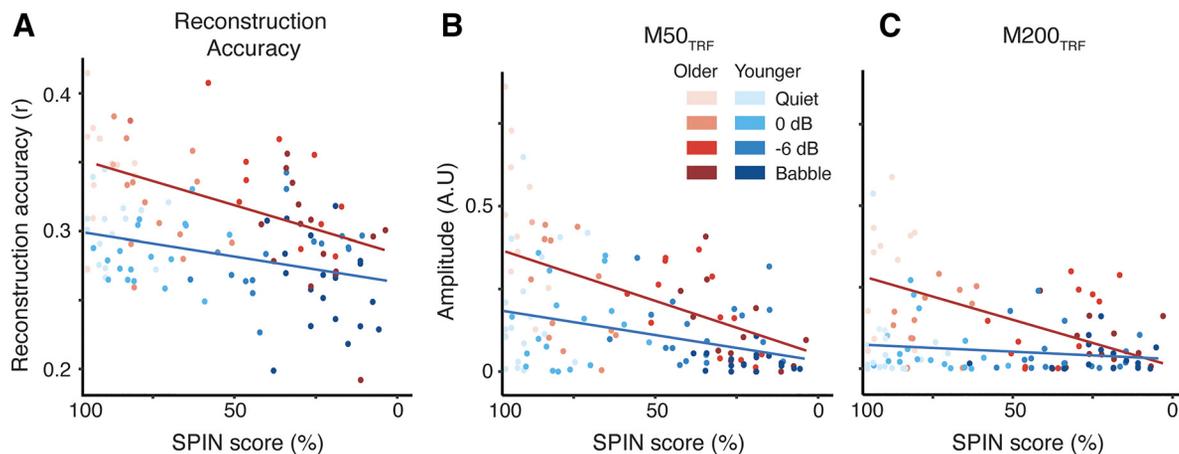


Figure 8. Neural measure vs. speech-in-noise (SPIN) scores including all noise conditions. A: reconstruction accuracy vs. SPIN score. B: temporal response function (TRF) peak with latency of ~ 50 ms ($M50_{TRF}$) peak amplitude vs. SPIN score. Better reconstruction accuracies or $M50_{TRF}$ peak amplitudes were related to better speech intelligibility scores in both groups. C: TRF peak with latency of ~ 200 ms ($M200_{TRF}$) peak amplitude vs. SPIN score. Only in older adults was a stronger $M200_{TRF}$ peak amplitude associated with better speech intelligibility score.

cortical processing stages associated with this exaggeration. Using the integration window analysis and TRF peaks, we have now shown that all major cortical processing stages, early, middle, and late processing, contributed to exaggerated neural responses. As previously shown, the addition of a competing talker diminishes the cortical representation of the attended speech signal, and here we have now also shown that aging enhances this reduction. In particular, TRF peak analysis has now revealed that it was only the middle and late processing contributions to the cortical representation that differ in amplitude between attended and unattended speech and that difference was affected by age and the interfering speech in a complex manner. Additionally, TRF peak latency analysis has now shown that all processing stages were delayed with interfering speech, which was further affected by aging. Details of these novel findings are addressed below.

Aging Is Associated with Exaggerated Speech Envelope Representation/Encoding

Perhaps counterintuitively, the reconstruction analysis demonstrated that compared with younger adults older adults exhibit exaggerated speech envelope representation irrespective of the noise condition. This replicates previous results by Presacco et al. (60) showing that older adults have a more robust (overly large) representation of the attended speech envelope in the cortex and is consistent with studies showing enhanced envelope tracking with advancing age, both for discrete stimuli (23, 61, 62) and continuous speech stimuli (13, 63). Whereas both Presacco et al. (60) and Decruy et al. (63) analyzed the attended speech envelope reconstruction, the present study extends the analysis to the unattended speech envelope reconstruction. Incorporating both attended and unattended talker representations allows investigation of the two speech streams as distinct auditory objects (64), separable via neural implementations of auditory scene analysis (65, 66). Here we have demonstrated that even the unattended speech envelope is exaggerated in the cortex of older adults and cortical exaggeration is not limited only to the attended speech. Moreover, the dynamical difference comparisons between age groups show that age-related exaggeration begins at latencies as early as 50–100 ms and continues as late as 350 ms. This suggests that neural mechanisms underlying the exaggerated representation are active even in the earliest cortical stages and some persist throughout the late processing stages.

The exaggerated envelope representation in older adults manifests as exaggerated TRF peak amplitudes at all three processing stages, $M50_{\text{TRF}}$, $M100_{\text{TRF}}$, and $M200_{\text{TRF}}$, and for both attended and unattended speech. The enhanced $M50_{\text{TRF}}$ in older adults also agrees with the integration window results above, that the exaggerated representation starts even at early cortical processing stages. Larger early cortical peaks (~50-ms latency) in older adults have been reported with both EEG (22, 67) and MEG (9, 36) for speech both in quiet and in noisy conditions. Alain et al. (68) suggested that this increased neural activity may be caused by excitatory/inhibitory imbalance, which is further in agreement with animal studies (22) and is consistent with other

studies (9). Larger cortical peaks at ~100-ms latency (e.g., the $M100_{\text{TRF}}$) in older adults have been reported with both EEG (22) and MEG (36), and the exaggeration is bigger for attended speech compared with unattended. The exaggerated response at this middle processing stage has been associated with increased task-related effort (69). However, previous studies have reported that the $M100$ is enhanced in older adults for both active and passive listening paradigms and also for simple stimuli (21, 70–72). This suggests that not just cognition per se but also other age-related functional changes may contribute to this enhancement, as elaborated in *Possible Mechanisms Underlying Exaggerated Speech Representation*. For longer latency cortical peaks (~200-ms latency, e.g., the $M200_{\text{TRF}}$), previous studies have reported an enhanced late peak in both EEG (73, 74) and MEG (36) when the stimulus was continuous speech. No age-related enhancement was seen for this stage, however, for a gap-in-noise detection task (29, 75). This may indicate that the late processing stage entails an additional stage of processing during speech comprehension that is not activated during simpler tasks such as tone processing.

Possible Mechanisms Underlying Exaggerated Speech Representation

Several potential mechanisms have been put forward to explain such exaggerated neural responses in older adults; not all of them necessarily apply for each of the three (early, middle, and late) processing stages, for both attended and unattended talkers. One well-supported explanation is an imbalance between excitatory and inhibitory currents, where a reduction in inhibition would result in greater neural currents and their electromagnetic fields but, because of the importance of inhibition for neuronal tuning, likely worse sensitivity, both temporally and spectrally (76). Gamma-aminobutyric acid (GABA)-mediated inhibition plays a major role in maintaining synchrony and spectral sensitivity in auditory circuits. Both animal (15, 16, 77–81) and human (82–85) studies have reported a reduction in age-related inhibitory circuits and function. This mechanism could apply to any of the three main processing stages and for both attended and unattended talkers. Another possible contributor to the age-related amplitude increase is additional auditory processing due to age-related reduction in cortical connectivity: Peelle et al. (86) found reduced coherence among cortical regions necessary to support speech comprehension, thus requiring multiple cortical regions to redundantly neurally process the same stimulus information, which would also result in increased extracranial neural responses. This top-down effect might be especially important for the middle and later processing stages. Finally, additional processing of the attended speech might arise from explicitly top-down compensatory mechanisms, where additional cortical regions would be recruited to support the early processing deficits (87–89). Imaging studies have shown that older adults, even in the absence of elevated hearing thresholds, engage more extensive and diffuse regions of frontal cortex at relatively lower task loads than younger adults (for a review, see Ref. 17). Linking age-related changes in neural activity to listening performance is critical for understanding the extent to which observed

upregulation of activity is evidence of a compensatory process (generally predictive of better listening performance) or of an inability to inhibit irrelevant cortical processing (i.e., dedifferentiation, predictive of poor performance; for a review, see Ref. 90)

Selective Attentional Modulation of Speech Representation/Encoding and Aging

In line with the results for younger adults (32, 74, 91, 92), we have found that in older adults the attended speech envelope is better represented than the unattended. Surprisingly, irrespective of exaggerated envelope representation in the cortex, we have shown that both groups showed similar effects of selective attention on the envelope representation, as no age \times selective attention interaction effect was found.

Both integration window analysis and TRF analysis contrasting attended versus unattended speech have here revealed that unattended speech is represented to a similar degree as attended (or even more strongly in the case of -6 dB SNR, when it is acoustically louder) at the early processing stage. However, by the middle and late processing stages, attended speech is better represented than unattended in both groups. Therefore, the early processing stages better reflect the full acoustic sound scene than the selective-attention-driven percept, whereas the middle and later stages more closely follow the percept (although see Ref. 93). Older adults exhibit an enhanced attended-unattended M100_{TRF} amplitude difference compared with younger adults (in addition to showing enhancement in both separately), which may reflect task-related increased attention or cognitive effort that further supports the selective attention.

Representations of Quiet and Noisy Speech Are Differentially Affected by Age

Our analysis also investigated how competing speakers affect the cortical speech representation. We found that, irrespective of age, masking by other speakers adversely affects the cortical representation of the attended speech envelope. This is also in line with previous studies that report that envelope tracking is adversely affected by the SNR (14, 33, 60, 91). Our results additionally show that older adults' envelope representation is more strongly affected, still negatively, by adding higher levels of noise, compared with younger adults. Thus, within an individual, better reconstruction accuracy is associated with clearer speech, but not across age groups; for example, older adults' reconstruction accuracy for 0 dB SNR is typically better than younger adults' reconstruction accuracy for quiet (as seen in Fig. 3A).

A drop in M50_{TRF} amplitude from quiet to noisy speech, and with similar TRF peak amplitudes for both attended and unattended speech at 0 dB SNR, is expected for an early auditory cortical stage that processes the complete acoustic scene and thus both attended and unattended talkers (73, 93). The older adults here showed greater reduction in M50_{TRF} amplitude with the noise condition, suggesting that aging adversely affects the early-stage cortical processing in speech-in-noise conditions. In contrast, older adults display increasing M100_{TRF} amplitude with the noise condition, whereas no significant M100_{TRF} amplitude differences are seen between noise conditions in younger adults. These

results are consistent with some previous studies (32, 94, 95), but not all (22, 36, 96), where a dependence on masker noise condition is found in both groups. Mechanistically, an exaggerated M100_{TRF} amplitude may reflect an increase in task-related attention or cognitive effort (69, 96), and as such any conflicting trends may be due to task difference subtleties. In contrast, the M200_{TRF} amplitude also has a pronounced decrease from quiet speech to noisy speech in older adults (no such drop is observed in younger adults). The M200_{TRF} amplitude is also strongly enhanced by selective attention, and it has a sufficiently long latency to reflect top-down compensatory processing known to be important for older adults (97). One possibility for the noise-related decrease in older adults is that the observed M200_{TRF} actually reflects the sum of two sources with similar latencies but opposite polarities, where the first (positive) source is active regardless of whether the speech is noisy but the second (negative and slightly later) top-down compensatory source is invoked only under difficult listening conditions; this finding is also consistent with the association between decreased M200_{TRF} peak amplitudes and longer latencies in older adults. This late peak in older adults has also been reported as a potential biomarker for behavioral inhibition (36), although the present study did not find any such correlations between M200_{TRF} amplitude and behavioral measures. The different amplitude trends as a function of masker level, between the M100_{TRF} and M200_{TRF} for the two groups, indicate that the presence and level of the masker significantly contribute to middle and late processing in older adults.

Aging Is Associated with Earlier Early Processing and Prolonged Late Processing

The integration window analysis revealed that a long temporal integration window (at least ~ 300 ms) allows a robust speech reconstruction for both groups, which is consistent with early studies using only younger adults (33, 74). Specifically, Power et al. (28) and O'Sullivan et al. (74) reported that an interval of duration 170–250 ms is important for attention decoding with EEG, and processing at that latency may even be at the level of semantic analysis. The striking difference between the age groups, however, is that older adults need more time to better represent the speech envelope, as seen in Fig. 4C, as long as 350 ms (14). This provides support for the existence of late compensatory mechanisms to support the additional selective attention filtering process, required for the early-stage processing deficits or slowing of synchronous neural firing rate (98, 99), which is addressed explicitly in the TRF analysis discussed next.

TRF peak latencies indicate processing time needed to generate responses after the corresponding acoustic feature and so can be mapped to the speed of auditory processing. Both age groups showed significant noise-related delays in the M50_{TRF}, M100_{TRF}, and M200_{TRF} peak latencies, suggesting longer cortical processing associated with the addition (and level) of the masker (22). The latency results for babble speech were less clear. Latencies for the babble condition were delayed compared with quiet speech, yet no consistent trends were observed compared with two-talker conditions. The babble speech may be impossibly challenging for some listeners, who are more likely to disengage attention,

reducing top-down effects, but for other listeners its challenge may not exceed limits, enhancing top-down effects, thus confounding comparisons between listeners (17).

Compared with younger adults, older adults demonstrated relatively early M50_{TRF} peaks. This effect has been observed in some studies using both CAEP (67, 98) and MEG (9) but not others (68, 99). This finding is consistent with an excitation/inhibition imbalance favoring excitation compared with younger adults. Another possible explanation is that an M50_{TRF} followed immediately by an exaggerated M100_{TRF} of the opposite polarity would appear shortened because of an earlier cutoff imposed by the subsequent peak, with the artifactual side effect of shorter latency. The M100_{TRF} latency was comparable for both groups, which contrasts with the late peak M200_{TRF}, which was significantly delayed in older adults. These findings are in line with studies that recorded responses to speech syllables (57, 100), suggesting an age-related decrease in rate of transmission for auditory neurons contributing to P2. In contrast to the early peak, both middle and late peaks demonstrated further delayed peak latencies with noise condition in older adults, suggesting overreliance on the middle and late processing mechanisms in older adults to compensate for degraded afferent input (101). Taken together, age-related impaired processing of the auditory input at the early stages could affect the auditory scene representation at the late processing stages, by employing additional cortical regions and compensatory mechanisms at the later stage.

Our results also add additional supporting evidence that the attended speech signal requires longer processing times in later stages to discern the information in the attended stream in older listeners, possibly to recover the early processing deficits by many compensatory mechanisms. Fiedler et al. (73), using EEG, demonstrated that late cortical tracking of the unattended talker reflects neural selectivity in acoustically challenging conditions. They reported an early suppressed positive P2 in line with the present study and additional late negative N2 peak for the unattended talker, which appears around the same latency as attended P2. They argue that this late N2 of the unattended talker actively suppresses the distracting inputs.

Behavior

As expected, SPIN scores and intelligibility ratings decreased as noise condition increased in both groups, suggesting that noise condition negatively affects speech intelligibility and in turn increases task difficulty. However, no significant difference was found between younger and older groups in the SPIN scores, which was unexpected. The SPIN measure employed here, developed from the same materials used during the MEG recordings, has not been calibrated against more standard SPIN measures and so may not be able to distinguish the hearing complications that arise with aging. For this task, subjects listened to a very short narrative segments (with 4–7 key words) with no time limit, and the older subjects may have benefited not only from their extended vocabulary and language experience (97, 102, 103) but also from the lack of time demands. In addition, the range of scores obtained in both groups suggests that the task was more challenging than other standardized measures, especially as some of the younger listeners did not achieve 100% performance even in the quiet

condition. In contrast, established speech intelligibility tests such as the QuickSIN (43) are known to show such behavioral age-related auditory declines (14, 60, 104). Unfortunately, because of the COVID-19 pandemic, QuickSIN measures were obtained from only half of the subjects, which was not enough to incorporate into analysis. The subjective intelligibility ratings of older adults were, perhaps surprisingly, higher than those of younger adults. This finding is nevertheless consistent with earlier results showing that older adults tend to underestimate their hearing difficulties relative to younger adults (105, 106). However, it is unclear whether it is younger adults who overestimate or older adults who underestimate their hearing difficulties (or both) and whether their subjective judgments are differently influenced by intelligibility, contextual factors, or other variables (107).

The positive association between the behavioral SPIN scores and intelligibility ratings does suggest that models using (subjective) intelligibility ratings collected during the task of interest would give outcomes similar to those using (objective) SPIN scores collected in a separate task (108). However, the age effects revealed only in subjective intelligibility ratings suggest that the two measures reflect different aspects of speech intelligibility or different factors that affect performance.

Our results also revealed that within a subject the neural measures of reconstruction accuracy M50_{TRF} and TRF peak amplitude are correlated with speech intelligibility (though speech intelligibility cannot be disentangled here from the changes in noise condition). The positive association between intelligibility and larger neural measures in general is suggestive of a compensatory (vs. dedifferentiation) pattern of neural engagement, in which additional resources are brought online to maintain performance in challenging conditions (90). Age-related compensation may be particularly evident during later-stage processing, as the correlation between the M200_{TRF} peak amplitude and better speech intelligibility score was only observed for older adults. Unexpectedly, we did not find any consistent relationships between neural measures and behavioral performance measures within a noise condition. One possible reason, as mentioned above, is that these uncalibrated behavioral performance measures may not sufficiently capture the known age-related hearing difficulties and temporal processing deficits.

Audiometry Differences

In the present study, we investigated the age-related cortical temporal processing deficits by recruiting younger and older subjects who have clinically normal hearing [pure-tone thresholds ≤ 25 dB hearing level (HL) from 125 to 4,000 Hz]. However, it is difficult to completely eliminate the confounds of peripheral hearing loss that gradually occur with aging. Previous studies have reported that this peripheral hearing loss may contribute to some of the speech-in-noise understanding difficulties experienced by older adults (109, 110). The average of the pure-tone thresholds (PTA) showed that there is a positive correlation between the age and PTA (Pearson's $r = 0.87$, $P < 0.001$). Therefore, the causal relationship between the hearing sensitivity and aging may cast doubt on the age-related neural changes we observed in the

present study. To answer this issue, we investigated the effects of hearing sensitivity (PTA) on the neural and behavioral measures analyzed in the present study. As there is a significant correlation between age and PTA, the LMER model wipes out the PTA from the models. Therefore, we evaluated the effects of PTA within an age group separately. However, we did not find any significant correlation between PTA and neural measures (reconstruction accuracy, TRF amplitudes and latencies) or between PTA and behavioral measures (intelligibility ratings and SPIN scores).

Age-related deterioration both peripherally, such as sub-clinical loss of outer and inner hair cells and ganglion cells within the cochlea (111, 112), and centrally, such as loss of neural synchrony (19, 25, 113, 114), may not affect audiometric thresholds but likely contributes to suprathreshold listening difficulties (115). Previous studies including both hearing-impaired and normal-hearing older adults have shown that beyond hearing sensitivity aging is the main driving factor of temporal processing differences in sub-cortical and cortical regions (100, 116). Therefore, these results suggest that aging and age-related hearing difficulties other than hearing sensitivity may cause the observed temporal processing differences in the auditory cortex. However, we acknowledge that audiometric thresholds may not adequately characterize peripheral auditory function and that the inclusion of other measures of peripheral function in the analysis, such as otoacoustic emissions or auditory brain stem response wave I amplitude, may reveal central consequences of decreased afferent input (117).

Conclusions

In conclusion, the present study showed that aging is associated with exaggerated speech representation in the cortical response and this exaggeration is noted in all three cortical processing stages; early, middle, and late processing. Moreover, the effects of speech intelligibility and attention on the respective three $M50_{TRF}$, $M100_{TRF}$, and $M200_{TRF}$ peak amplitudes and latencies reveal characteristics related to different cortical processing stages, and aging appears to differently affect the individual processing stages. Earlier and enhanced processing of early stages supports the hypothesis of an excitatory/inhibitory imbalance in older adults. The underlying causes of enhanced processing of the middle cortical processing stage are less clear; indeed, the strongest enhancement occurs under noise conditions in which the timing is neither earlier nor later than for younger adults. The late auditory cortical stage displays both delayed and enhanced processing in older adults, which were associated with better performance on a parallel speech-in-noise task and thus consistent with cognitive (perhaps memory/experience based) compensatory mechanisms. Overall, these findings support the theory that some of the age-related difficulties in understanding speech in noise experienced by older adults, including those directly related to temporal processing, are accompanied by age-related temporal processing differences in auditory cortex.

DATA AVAILABILITY

Preprocessed MEG data, stimulus material, behavioral data, and analysis codes are available at <https://doi.org/10.13016/4goq-vs0z>.

SUPPLEMENTAL MATERIALS

Supplemental Tables S1–S10: <https://doi.org/10.13016/4goq-vs0z>.

ACKNOWLEDGMENTS

We thank Katie Brow, Julie Cohen, and Carol Gorham for help in recruiting and testing subjects.

GRANTS

This work was supported by grants from the National Institute on Aging (P01-AG055365) and the National Institute of Deafness and Other Communication Disorders (R01-DC014085, R01-DC019394).

DISCLAIMERS

The identification of specific products or scientific instrumentation is considered an integral part of the scientific endeavor and does not constitute endorsement or implied endorsement on the part of the authors, DoD, or any component agency. The views expressed in this paper are those of the authors and do not necessarily reflect the official policy of the Department of Defense or the US Government.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

I.M.D.K., J.L.D., J.P., A.P., S.A., S.E.K., and J.Z.S. conceived and designed the research; I.M.D.K., J.L.D., J.P., and L.D. performed experiments; I.M.D.K., J.L.D., L.D., and J.Z.S. analyzed data; I.M.D.K., L.D., S.A., S.E.K., and J.Z.S. interpreted results of experiments; I.M.D.K. prepared figures; I.M.D.K. drafted manuscript; I.M.D.K., J.L.D., J.P., A.P., L.D., S.A., S.E.K., and J.Z.S. edited and revised manuscript; I.M.D.K., J.L.D., J.P., A.P., L.D., S.A., S.E.K., and J.Z.S. approved final version of manuscript.

REFERENCES

1. **Fitzgibbons PJ, Gordon-Salant S.** Age effects on duration discrimination with simple and complex stimuli. *J Acoust Soc Am* 98: 3140–3145, 1995. doi:10.1121/1.413803.
2. **Snell KB.** Age-related changes in temporal gap detection. *J Acoust Soc Am* 101: 2214–2220, 1997. doi:10.1121/1.418205.
3. **Frisina DR, Frisina RD.** Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res* 106: 95–104, 1997. doi:10.1016/S0378-5955(97)00006-3.
4. **Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Barrett J.** Age-related differences in identification and discrimination of temporal cues in speech segments. *J Acoust Soc Am* 119: 2455–2466, 2006. doi:10.1121/1.2171527.
5. **Bess FH, Lichtenstein MJ, Logan SA, Burger MC, Nelson E.** Hearing impairment as a determinant of function in the elderly. *J Am Geriatr Soc* 37: 123–128, 1989. doi:10.1111/j.1532-5415.1989.tb05870.x.
6. **Strawbridge WJ, Wallhagen MI, Shema SJ, Kaplan GA.** Negative consequences of hearing impairment in old age. *Gerontologist* 40: 320–326, 2000. doi:10.1093/geront/40.3.320.
7. **Gopinath B, Wang JJ, Schneider J, Burlutsky G, Snowdon J, McMahon CM, Leeder SR, Mitchell P.** Depressive symptoms in older adults with hearing impairments: the Blue Mountains Study. *J Am Geriatr Soc* 57: 1306–1308, 2009. doi:10.1111/j.1532-5415.2009.02317.x.

8. Uhlmann RF, Larson EB, Rees TS, Koepsell TD, Duckert LG. Relationship of hearing impairment to dementia and cognitive dysfunction in older adults. *JAMA* 261: 1916–1919, 1989.
9. Brodbeck C, Presacco A, Anderson S, Simon JZ. Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust United Acust* 104: 774–777, 2018. doi:10.3813/AAA.919221.
10. Decruy L, Vanthornhout J, Francart T. Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *J Neurophysiol* 122: 601–615, 2019. doi:10.1152/jn.00687.2018.
11. Gillis M, Kries J, Vandermosten M, Francart T. Neural tracking of linguistic and acoustic speech representations decreases with advancing age. *NeuroImage* 267: 119841, 2023. doi:10.1016/j.neuroimage.2022.119841.
12. Manan HA, Yusoff AN, Franz EA, Mukari SM. Effects of aging and background babble noise on speech perception processing: an fMRI study. *Neurophysiology* 49: 441–452, 2017. doi:10.1007/s11062-018-9707-5.
13. Mesik J, Ray L, Wojtczak M. Effects of age on cortical tracking of word-level features of continuous competing speech. *Front Neurosci* 15: 635126, 2021. doi:10.3389/fnins.2021.635126.
14. Presacco A, Simon JZ, Anderson S. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol* 116: 2346–2355, 2016. doi:10.1152/jn.00372.2016.
15. Caspary DM, Hughes LF, Ling LL. Age-related GABA_A receptor changes in rat auditory cortex. *Neurobiol Aging* 34: 1486–1496, 2013. doi:10.1016/j.neurobiolaging.2012.11.009.
16. Hughes LF, Turner JG, Parrish JL, Caspary DM. Processing of broadband stimuli across A1 layers in young and aged rats. *Hear Res* 264: 79–85, 2010. doi:10.1016/j.heares.2009.09.005.
17. Kuchinsky SE, Vaden KI. Aging, hearing loss, and listening effort: imaging studies of the aging listener. In: *Aging and Hearing*, edited by Helfer KS, Bartlett EL, Popper AN, Fay RR. Cham, Switzerland: Springer International Publishing, 2020, p. 231–256.
18. Dryden A, Allen HA, Henshaw H, Heinrich A. The association between cognitive performance and speech-in-noise perception for adult listeners: a systematic literature review and meta-analysis. *Trends Hear* 21: 233121651774467, 2017. doi:10.1177/2331216517744675.
19. Schneider BA, Pichora-Fuller MK. Implications of perceptual deterioration for cognitive aging research. In: *The Handbook of Aging and Cognition*, edited by Craik FI, Salthouse TA. Mahwah, NJ: Lawrence Erlbaum Associates, 2000, p. 755.
20. Tremblay KL, Piskosz M, Souza P. Aging alters the neural representation of speech cues. *NeuroReport* 13: 1865–1870, 2002. doi:10.1097/00001756-200210280-00007.
21. Tremblay KL, Piskosz M, Souza P. Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol* 114: 1332–1343, 2003. doi:10.1016/s1388-2457(03)00114-7.
22. McCullagh J, Shinn JB. Auditory cortical processing in noise in younger and older adults. *Hear Balance Commun* 11: 182–190, 2013. doi:10.3109/21695717.2013.855374.
23. Bidelman GM, Villafuerte JW, Moreno S, Alain C. Age-related changes in the subcortical–cortical encoding and categorical perception of speech. *Neurobiol Aging* 35: 2526–2540, 2014. doi:10.1016/j.neurobiolaging.2014.05.006.
24. Tremblay K, Ross B. Effects of age and age-related hearing loss on the brain. *J Commun Disord* 40: 305–312, 2007. doi:10.1016/j.jcomdis.2007.03.008.
25. Anderson S, Karawani H. Objective evidence of temporal processing deficits in older adults. *Hear Res* 397: 108053, 2020. doi:10.1016/j.heares.2020.108053.
26. Keidser G, Naylor G, Brungart DS, Caduff A, Campos J, Carlile S, Carpenter MG, Grimm G, Hohmann V, Holube I, Launer S, Lunner T, Mehra R, Rapport F, Slaney M, Smeds K. The quest for ecological validity in hearing science: what it is, why it matters, and how to advance it. *Ear Hear* 41: 5S–19S, 2020. doi:10.1097/AUD.0000000000000944.
27. Ding N, Simon JZ. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol* 107: 78–89, 2012. doi:10.1152/jn.00297.2011.
28. Power AJ, Foxe JJ, Forde E-J, Reilly RB, Lalor EC. At what time is the cocktail party? A late locus of selective attention to natural speech. *Eur J Neurosci* 35: 1497–1503, 2012. doi:10.1111/j.1460-9568.2012.08060.x.
29. Lister JJ, Maxfield ND, Pitt GJ, Gonzalez VB. Auditory evoked response to gaps in noise: older adults. *Int J Audiol* 50: 211–225, 2011. doi:10.3109/14992027.2010.526967.
30. Ceponiene R, Alku P, Westerfield M, Torki M, Townsend J. ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology* 42: 391–406, 2005. doi:10.1111/j.1469-8986.2005.00305.x.
31. Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull* 125: 826–859, 1999. doi:10.1037/0033-2909.125.6.826.
32. Ding N, Simon JZ. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc Natl Acad Sci USA* 109: 11854–11859, 2012. doi:10.1073/pnas.1205381109.
33. Ding N, Simon JZ. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J Neurosci* 33: 5728–5735, 2013. doi:10.1523/JNEUROSCI.5297-12.2013.
34. Sheehan KA, McArthur GM, Bishop DV. Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Brain Res Cogn Brain Res* 25: 547–553, 2005. doi:10.1016/j.cogbrainres.2005.08.007.
35. Tong Y, Melara RD, Rao A. P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Res* 1297: 80–88, 2009. doi:10.1016/j.brainres.2009.07.089.
36. Zan P, Presacco A, Anderson S, Simon JZ. Exaggerated cortical representation of speech in older listeners: mutual information analysis. *J Neurophysiol* 124: 1152–1164, 2020. doi:10.1152/jn.00002.2020.
37. de Cheveigné A, Simon JZ. Denoising based on time-shift PCA. *J Neurosci Methods* 165: 297–305, 2007. doi:10.1016/j.jneumeth.2007.06.003.
38. de Cheveigné A, Simon JZ. Sensor noise suppression. *J Neurosci Methods* 168: 195–202, 2008. doi:10.1016/j.jneumeth.2007.09.012.
39. de Cheveigné A, Simon JZ. Denoising based on spatial filtering. *J Neurosci Methods* 171: 331–339, 2008. doi:10.1016/j.jneumeth.2008.03.015.
40. Särelä J, Valpola H. Denoising source separation. *J Mach Learn Res* 6: 233–272, 2005.
41. Biesmans W, Das N, Francart T, Bertrand A. Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Trans Neural Syst Rehabil Eng* 25: 402–412, 2017. doi:10.1109/TNSRE.2016.2571900.
42. Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ. NIH Toolbox for Assessment of Neurological and Behavioral Function. *Neurology* 80: S2–S6, 2013. doi:10.1212/WNL.0b013e3182872e5f.
43. Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116: 2395–2405, 2004 [Erratum in *J Acoust Soc Am* 119: 1888, 2006]. doi:10.1121/1.1784440.
44. David SV, Mesgarani N, Shamma SA. Estimating sparse spectrotemporal receptive fields with natural stimuli. *Network* 18: 191–212, 2007. doi:10.1080/09548980701609235.
45. Ding N, Chatterjee M, Simon JZ. Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage* 88: 41–46, 2014. doi:10.1016/j.neuroimage.2013.10.054.
46. Haufe S, Meinecke F, Görgen K, Dähne S, Haynes JD, Blankertz B, Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87: 96–110, 2014. doi:10.1016/j.neuroimage.2013.10.067.
47. Lalor EC, Power AJ, Reilly RB, Foxe JJ. Resolving precise temporal processing properties of the auditory system using continuous stimuli. *J Neurophysiol* 102: 349–359, 2009. doi:10.1152/jn.90896.2008.
48. Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE. Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* 77: 980–991, 2013. doi:10.1016/j.neuron.2012.12.037.
49. R Core Team. R: A Language and Environment for Statistical Computing (Online). R Foundation for Statistical Computing. 2020. <https://www.R-project.org/>.

50. **Bates D, Mächler M, Bolker B, Walker S.** Fitting linear mixed-effects models using lme4. *J Stat Soft* 67: 1–48, 2015. doi:10.18637/jss.v067.i01.
51. **Kuznetsova A, Brockhoff PB, Christensen RH.** lmerTest package: tests in linear mixed effects models. *J Stat Soft* 82: 1–26, 2017. doi:10.18637/jss.v082.i13.
52. **Voeten CC.** buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression (Online). 2020. <https://CRAN.R-project.org/package=buildmer>.
53. **Luke SG.** Evaluating significance in linear mixed-effects models in R. *Behav Res Methods* 49: 1494–1502, 2017. doi:10.3758/s13428-016-0809-y.
54. **Satterthwaite FE.** Synthesis of variance. *Psychometrika* 6: 309–316, 1941. doi:10.1007/BF02288586.
55. **Wood SN.** *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman & Hall/CRC, 2006.
56. **DeCat C, Baayen H, Klepousniotou E.** Electrophysiological correlates of noun-noun compound processing by non-native speakers of English. In *Proceedings of the First Workshop on Computational Approaches to Compound Analysis (ComAComA)*. 2014. Association for Computational Linguistics and Dublin City University, p. 41–52.
57. **Tremblay A, Newman AJ.** Modeling nonlinear relationships in ERP data using mixed-effects regression with R examples: modeling using mixed-effects regression. *Psychophysiology* 52: 124–139, 2015. doi:10.1111/psyp.12299.
58. **van Rij J, Hendriks P, van Rijn H, Baayen RH, Wood SN.** Analyzing the time course of pupillometric data. *Trends Hear* 23: 233121651983248, 2019. doi:10.1177/2331216519832483.
59. **Sóskuthy M.** Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction (Preprint). *ArXiv* 1703.05339, 2017. doi:10.48550/arXiv.1703.05339.
60. **Presacco A, Simon JZ, Anderson S.** Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol* 116: 2356–2367, 2016. doi:10.1152/jn.00373.2016.
61. **Goossens T, Vercammen C, Wouters J, Wieringen A. V.** Aging affects neural synchronization to speech-related acoustic modulations. *Front Aging Neurosci* 8: 133, 2016. doi:10.3389/fnagi.2016.00133.
62. **Irsik VC, Almanaseer A, Johnsrude IS, Herrmann B.** Cortical responses to the amplitude envelopes of sounds change with age. *J Neurosci* 41: 5045–5055, 2021. doi:10.1523/JNEUROSCI.2715-20.2021.
63. **Decruy L, Vanthornhout J, Francart T.** Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hear Res* 393: 107961, 2020. doi:10.1016/j.heares.2020.107961.
64. **Griffiths TD, Warren JD.** What is an auditory object? *Nat Rev Neurosci* 5: 887–892, 2004. doi:10.1038/nrn1538.
65. **Shamma SA, Elhilali M, Micheyl C.** Temporal coherence and attention in auditory scene analysis. *Trends Neurosci* 34: 114–123, 2011. doi:10.1016/j.tins.2010.11.002.
66. **Shinn-Cunningham BG.** Object-based auditory and visual attention. *Trends Cogn Sci* 12: 182–186, 2008. doi:10.1016/j.tics.2008.02.003.
67. **Roque L, Karawani H, Gordon-Salant S, Anderson S.** Effects of age, cognition, and neural encoding on the perception of temporal speech cues. *Front Neurosci* 13: 749, 2019. doi:10.3389/fnins.2019.00749.
68. **Alain C, Roye A, Salloum C.** Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. *Front Syst Neurosci* 8: 8, 2014. doi:10.3389/fnsys.2014.00008.
69. **Rao A, Zhang Y, Miller S.** Selective listening of concurrent auditory stimuli: an event-related potential study. *Hear Res* 268: 123–132, 2010. doi:10.1016/j.heares.2010.05.013.
70. **Herrmann B, Maess B, Johnsrude IS.** Aging affects adaptation to sound-level statistics in human auditory cortex. *J Neurosci* 38: 1989–1999, 2018. doi:10.1523/JNEUROSCI.1489-17.2018.
71. **Herrmann B, Maess B, Johnsrude IS.** A neural signature of regularity in sound is reduced in older adults. *Neurobiol Aging* 109: 1–10, 2022. doi:10.1016/j.neurobiolaging.2021.09.011.
72. **Sörös P, Teismann IK, Manemann E, Lütkenhöner B.** Auditory temporal processing in healthy aging: a magnetoencephalographic study. *BMC Neurosci* 10: 34, 2009. doi:10.1186/1471-2202-10-34.
73. **Fiedler L, Wöstmann M, Herbst SK, Obleser J.** Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *NeuroImage* 186: 33–42, 2019. doi:10.1016/j.neuroimage.2018.10.057.
74. **O’Sullivan JA, Power AJ, Mesgarani N, Rajaram S, Foxe JJ, Shinn-Cunningham BG, Slaney M, Shamma SA, Lalor EC.** Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb Cortex* 25: 1697–1706, 2015. doi:10.1093/cercor/bht355.
75. **Alain C, Snyder JS.** Age-related differences in auditory evoked responses during rapid perceptual learning. *Clin Neurophysiol* 119: 356–366, 2008. doi:10.1016/j.clinph.2007.10.024.
76. **Herrmann B, Butler BE.** Hearing loss and brain plasticity: the hyperactivity phenomenon. *Brain Struct Funct* 226: 2019–2039, 2021. doi:10.1007/s00429-021-02313-9.
77. **Caspary DM, Milbrandt JC, Helfert RH.** Central auditory aging: GABA changes in the inferior colliculus. *Exp Gerontol* 30: 349–360, 1995. doi:10.1016/0531-5565(94)00052-5.
78. **Parthasarathy A, Herrmann B, Bartlett EL.** Aging alters envelope representations of speech-like sounds in the inferior colliculus. *Neurobiol Aging* 73: 30–40, 2019. doi:10.1016/j.neurobiolaging.2018.08.023.
79. **Ramamurthy DL, Recanzone GH.** Age-related changes in sound onset and offset intensity coding in auditory cortical fields A1 and CL of rhesus macaques. *J Neurophysiol* 123: 1015–1025, 2020. doi:10.1152/jn.00373.2019.
80. **Richardson BD, Ling LL, Uteshev VV, Caspary DM.** Reduced GABA_A receptor-mediated tonic inhibition in aged rat auditory thalamus. *J Neurosci* 33: 1218–1227a, 2013. doi:10.1523/JNEUROSCI.3277-12.2013.
81. **Willott JF, Parham K, Hunter KP.** Comparison of the auditory sensitivity of neurons in the cochlear nucleus and inferior colliculus of young and aging C57BL/6J and CBA/J mice. *Hear Res* 53: 78–94, 1991. doi:10.1016/0378-5955(91)90215-u.
82. **Dobri SG, Ross B.** Total GABA level in human auditory cortex is associated with speech-in-noise understanding in older age. *NeuroImage* 225: 117474, 2021. doi:10.1016/j.neuroimage.2020.117474.
83. **Harris KC, Dias JW, McClaskey CM, Rumschlag J, Prisciandaro J, Dubno JR.** Afferent loss, GABA, and central gain in older adults: associations with speech recognition in noise. *J Neurosci* 42: 7201–7212, 2022. doi:10.1523/JNEUROSCI.0242-22.2022.
84. **Lalwani P, Gagnon H, Cassady K, Simmonite M, Peltier S, Seidler RD, Taylor SF, Weissman DH, Polk TA.** Neural distinctiveness declines with age in auditory cortex and is associated with auditory GABA levels. *NeuroImage* 201: 116033, 2019. doi:10.1016/j.neuroimage.2019.116033.
85. **Ross B, Dobri S, Schumann A.** Speech-in-noise understanding in older age: the role of inhibitory cortical responses. *Eur J Neurosci* 51: 891–908, 2020. doi:10.1111/ejn.14573.
86. **Peelle JE, Troiani V, Wingfield A, Grossman M.** Neural processing during older adults’ comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cereb Cortex* 20: 773–782, 2010. doi:10.1093/cercor/bhp142.
87. **Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BW, Humes LE, Lemke U, Lunner T, Matthen M, Mackersie CL, Naylor G, Phillips NA, Richter M, Rudner M, Sommers MS, Tremblay KL, Wingfield A.** Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear* 37: 5S–27S, 2016. doi:10.1097/AUD.0000000000000312.
88. **Rumschlag JA, McClaskey CM, Dias JW, Kerouac LB, Noble KV, Panganiban C, Lang H, Harris KC.** Age-related central gain with degraded neural synchrony in the auditory brainstem of mice and humans. *Neurobiol Aging* 115: 50–59, 2022. doi:10.1016/j.neurobiolaging.2022.03.014.
89. **Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS.** Effortful listening: the processing of degraded speech depends critically on attention. *J Neurosci* 32: 14010–14021, 2012. doi:10.1523/JNEUROSCI.1528-12.2012.
90. **Wingfield A, Grossman M.** Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. *J Neurophysiol* 96: 2830–2839, 2006. doi:10.1152/jn.00628.2006.
91. **Das N, Bertrand A, Francart T.** EEG-based auditory attention detection: boundary conditions for background noise and speaker positions. *J Neural Eng* 15: 066017, 2018. doi:10.1088/1741-2552/aae0a6.

92. **Mesgarani N, Chang EF.** Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485: 233–236, 2012. doi:10.1038/nature11020.
93. **Brodbeck C, Jiao A, Hong LE, Simon JZ.** Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLOS Biol* 18: e3000883, 2020. doi:10.1371/journal.pbio.3000883.
94. **McCullagh J, Musiek FE, Shinn JB.** Auditory cortical processing in noise in normal-hearing young adults. *Audiol Med* 10: 114–121, 2012. doi:10.3109/1651386X.2012.707354.
95. **Rufener KS, Liem F, Meyer M.** Age-related differences in auditory evoked potentials as a function of task modulation during speech–nonspeech processing. *Brain Behav* 4: 21–28, 2014. doi:10.1002/brb3.188.
96. **Billings CJ, Penman TM, McMillan GP, Ellis EM.** Electrophysiology and perception of speech in noise in older listeners: effects of hearing impairment and age. *Ear Hear* 36: 710–722, 2015. doi:10.1097/AUD.0000000000000191.
97. **Pichora-Fuller MK.** Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int J Audiol* 47: S72–S82, 2008. doi:10.1080/14992020802307404.
98. **O'Brien JL, Nikjeh DA, Lister JJ.** Interaction of musicianship and aging: a comparison of cortical auditory evoked potentials. *Behav Neurol* 2015: 545917, 2015. doi:10.1155/2015/545917.
99. **Tremblay KL, Billings C, Rohila N.** Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 15: 226–237, 2004. doi:10.3766/jaaa.15.3.5.
100. **Roque L, Gaskins C, Gordon-Salant S, Goupell MJ, Anderson S.** Age effects on neural representation and perception of silence duration cues in speech. *J Speech Lang Hear Res* 62: 1099–1116, 2019. doi:10.1044/2018_JSLHR-H-ASCC7-18-0076.
101. **Parthasarathy A, Hancock KE, Bennett K, DeGruttola V, Polley DB.** Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *eLife* 9: e51419, 2020. doi:10.7554/eLife.51419.
102. **Pichora-Fuller MK, Schneider BA, Daneman M.** How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 97: 593–608, 1995. doi:10.1121/1.412282.
103. **Schneider BA, Avivi-Reich M, Leung C, Heinrich A.** How age and linguistic competence affect memory for heard information. *Front Psychol* 7: 618, 2016. doi:10.3389/fpsyg.2016.00618.
104. **Holder JT, Levin LM, Gifford RH.** Speech recognition in noise for adults with normal hearing: age-normative performance for AzBio, BKB-SIN, and QuickSIN. *Otol Neurotol* 39: e972–e978, 2018. doi:10.1097/MAO.0000000000002003.
105. **Gosselin PA, Gagné JP.** Older adults expend more listening effort than young adults recognizing speech in noise. *J Speech Lang Hear Res* 54: 944–958, 2011. doi:10.1044/1092-4388(2010/10-0069).
106. **Uchida Y, Nakashima T, Ando F, Niino N, Shimokata H.** Prevalence of self-perceived auditory problems and their relation to audiometric thresholds in a middle-aged to elderly population. *Acta Otolaryngol* 123: 618–626, 2003. doi:10.1080/00016480310001448.
107. **Perron M, Dimitrijevic A, Alain C.** Objective and subjective hearing difficulties are associated with lower inhibitory control. *Ear Hear* 43: 1904–1916, 2022. doi:10.1097/AUD.0000000000001227.
108. **Hazan V, Tuomainen O, Tu L, Kim J, Davis C, Brungart D, Sheffield B.** How do aging and age-related hearing loss affect the ability to communicate effectively in challenging communicative conditions? *Hear Res* 369: 33–41, 2018. doi:10.1016/j.heares.2018.06.009.
109. **Motlagh Zadeh L, Silbert NH, Sternasty K, Swanepoel DW, Hunter LL, Moore DR.** Extended high-frequency hearing enhances speech perception in noise. *Proc Natl Acad Sci USA* 116: 23753–23759, 2019. doi:10.1073/pnas.1903315116.
110. **Yeend I, Beach EF, Sharma M.** Working memory and extended high-frequency hearing in adults: diagnostic predictors of speech-in-noise perception. *Ear Hear* 40: 458–467, 2019. doi:10.1097/AUD.0000000000000640.
111. **Parthasarathy A, Kujawa SG.** Synaptopathy in the aging cochlea: characterizing early-neural deficits in auditory temporal envelope processing. *J Neurosci* 38: 7108–7119, 2018. doi:10.1523/JNEUROSCI.3240-17.2018.
112. **Wu PZ, Liberman LD, Bennett K, de Gruttola V, O'Malley JT, Liberman MC.** Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience* 407: 8–20, 2019. doi:10.1016/j.neuroscience.2018.07.053.
113. **Boettcher FA, Mills JH, Norton BL.** Age-related changes in auditory evoked potentials of gerbils. I. Response amplitudes. *Hear Res* 71: 137–145, 1993. doi:10.1016/0378-5955(93)90029-z.
114. **Chis olm TH, Willott JF, Lister JJ.** The aging auditory system: anatomic and physiologic changes and implications for rehabilitation. *Int J Audiol* 42: 3–10, 2003. doi:10.3109/14992020309074637.
115. **Plack CJ, Barker D, Prendergast G.** Perceptual consequences of “hidden” hearing loss. *Trends Hear* 18: 233121651455062, 2014. doi:10.1177/2331216514550621.
116. **Presacco A, Simon JZ, Anderson S.** Speech-in-noise representation in the aging midbrain and cortex: effects of hearing loss. *PLoS One* 14: e0213899, 2019. doi:10.1371/journal.pone.0213899.
117. **Anderson S, Bieber R, Schloss A.** Peripheral deficits and phase-locking declines in aging adults. *Hear Res* 403: 108188, 2021. doi:10.1016/j.heares.2021.108188.