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], Jonathan Z. Simon[1,3,5]

[1]*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA*

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

[3]*Institute for Systems Research, University of Maryland, College Park, MD 20742, USA*

[4]*Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, Maryland, USA*

[5]*Department of Biology, University of Maryland, College Park, MD 20742, USA*

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 are 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 using 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 to 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.

Keywords

aging; electrophysiology; cortex; auditory; hearing

NEW & NOTEWORTHY

Older adults' difficulty understanding speech in noise may be related to age-related changes in cortical temporal processing. Using magnetoencephalography to record responses of listeners under different noise conditions, we investigated both timing and strength of the cortical representation of continuous speech at several cortical processing stages. The representation at each stage depends differently on noise level and selective attention, and in different ways for older listeners, even with normal hearing, than younger.

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. Behavioral studies have observed that some of these age-related changes in speech processing may be due to temporal processing deficits that occur in older adults, even those with hearing thresholds within clinically normal limits. Compared to 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 (Fitzgibbons and Gordon□Salant 1995) and gap detection (Snell 1997), or in acoustically complex paradigms such as speech listening in noise (Frisina and Frisina 1997; Gordon-Salant et al. 2006). Because poor speech comprehension in noise is associated with adverse psycho-social effects (Bess et al. 1989; Strawbridge et al. 2000), depression (Gopinath et al. 2009) and dementia (Uhlmann et al. 1989), 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. Previous studies have investigated age-related changes in the neural processing of phonemes, words, and phrases, but there is limited understanding of the extent to which the neural timing and fidelity of attended and unattended continuous speech streams contribute to the challenges that older adults face when listening to connected discourse.

Numerous studies have demonstrated important age-related anatomical and functional changes within peripheral and central auditory nervous systems that may contribute to listening difficulties. Animal and neurophysiological studies have reported age-related deterioration both peripherally, such as loss of outer and inner hair cells and ganglion cells within the cochlea (Parthasarathy and Kujawa 2018; Wu et al. 2019), and centrally, such as loss of neural synchrony (Boettcher et al. 1993; Schneider and Pichora-Fuller 2000; Chisolm et al. 2003; Anderson and Karawani 2020). This latter deterioration may not affect audiometric thresholds, but likely contributes to suprathreshold listening difficulties (Plack et al. 2014). Animal studies have also reported age-related excitatory and inhibitory imbalance in the ventral cochlear nucleus, dorsal cochlear nucleus, inferior colliculus and auditory cortex (Willott et al. 1991; Caspary et al. 1995; Hughes et al. 2010; Parthasarathy and Kujawa 2018), additionally leading to altered neural coding in the auditory pathway. In studies with human participants, abnormal neural activity patterns in the auditory cortex have been shown in older adults during speech-in-noise tasks (Presacco et al. 2016a, 2016b; Manan et al. 2017; Brodbeck et al. 2018; Decruy et al. 2019; Mesik et al. 2021), potentially reflecting changes in the central auditory system that could contribute to speech comprehension difficulties in older adults.

Beyond the detection and processing of low-level auditory information, speech processing also entails real-time recognition of phonemes, semantic decoding, and integration with long term memory, not only to perceive, but also to comprehend speech, particularly for connected discourse. Age-related changes in auditory, linguistic, and cognitive processes can influence the ease of speech understanding (for a review, see Kuchinsky and Vaden (2020). In noisy environments speech comprehension becomes even more challenging if aging is accompanied by deterioration of complementary executive functions, including attention (McDowd and Shaw 2000), working memory (Fabiani 2012) and processing speed (Eckert et al. 2010). Some studies have reported that age-related changes in cognitive functions limit successful speech understanding among older adults (Dryden et al. 2017). These findings highlight that age-related hearing difficulties are due to a complex combination of anatomical, functional and cognitive factors (Schneider and Pichora-Fuller 2000). Therefore, audiometric measurements of the peripheral auditory system alone may not be sufficient to evaluate and manage the hearing difficulties reported by older adults, and consideration of age and speech-in-noise recognition abilities may provide a better estimate of speech comprehension in everyday environments (Phatak et al. 2019). Measures of central auditory functions and related cognitive functions could be incorporated to both diagnostic evaluations and treatments aimed at speech comprehension problems in older adults.

The human neurophysiology underlying age-related changes in the timing and fidelity of encoding of connected speech is not well understood. Previous studies have examined neuromarkers of auditory encoding that include peaks in an auditory evoked neural response, where peak strength and latency may be tied to successful processing. High temporal resolution recording methods such as electroencephalography (EEG) or magnetoencephalography (MEG) are well-suited to accurately estimate such neural responses. Much of this research focuses on peaks in the auditory evoked response, obtained by averaging response to many repetitions of simple stimuli. Evoked response studies have reported decrease in the strength of the neural signal generated in subcortical areas with aging (Anderson et al. 2012; Bidelman et al. 2014; Anderson and Karawani 2020). 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 (Keidser et al. 2020).

Computational tools have been developed that can analyze neural responses to *continuous speech*, typically in terms of speech encoding and decoding models. Speech 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 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 (Presacco et al. 2016a, 2016b; Decruy et al. 2019), but little is known about the extent to which this may depend on the representation of the interfering speech, as a function of aging. Quantifying the timing and fidelity of unattended speech stream processing and how it affects attended speech processing analysis is critical to understanding how aging affects auditory scene representation in general, and stream segregation in particular, in the cortex.

The above-mentioned speech reconstruction analysis effectively utilizes the fine time resolution of MEG and EEG, but only when integrated over a longer time window (typically 500 ms), making estimating the latency time-course of age-related over-representation more difficult. Reconstruction accuracy values depend, in general, on the specifics of the temporal analysis window employed, and is thus indirectly affected by the latencies (cortical stages) at which any overrepresentation starts and ends. If aging affects the time course of speech processing, then the amount of usable speech information available in systematically varied reconstruction time windows should reflect changes in the length of time required to construct and maintain the continuous speech signal. Presacco et al. (2016a) 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. (2016a) using a nonlinear mixed effects modeling approach (i.e., generalized additive mixed effects models, GAMMs) to better understand the time course of speech processing.

Furthermore, while temporal integration window analysis is beneficial for understanding the evolution of the envelope representation from the early to late cortical processing stages, more detailed temporal information can be gained more directly from a TRF analysis. Therefore, in our study, we also investigated the impact of different evoked latencies on older adults' neural response using TRF analysis (Ding and Simon 2012a; Power et al. 2012). 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 (Lister et al. 2011). For evoked responses that may have analogous peaks at the corresponding latencies, it has been suggested that the early $M50_{TRF}$ peak dominantly reflects the neural encoding of acoustic features (Näätänen and Winkler 1999; Ceponiene et al. 2005), whereas the $M100_{TRF}$ peak reflects processing of selectively attended features (Näätänen and Winkler 1999). 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 dependence (Ding and Simon 2012b, 2013). The late peak $M200_{TRF}$ (~200 ms) has not previously been investigated in depth; it is quite late for encoding acoustic features, but appropriately positioned to reflect a representation of auditory object formation (Näätänen and Winkler 1999).

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 participants who had hearing thresholds within clinically normal limits through 4.0 kHz were recruited in the study.

METHODS

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

Participants

34 native English speakers participated in the study: 18 younger adults (7 males; mean age 20 y, range 17-26 y) and 16 older adults (5 males; mean age 70 y, range 65-78 y). Data from two additional subjects (1 older and 1 younger) were not included in the analysis due to data saturation caused by excessive MEG artifacts. All participants had normal hearing (see Figure 1), defined as pure-tone thresholds ≤ 25 dB hearing level (HL) from 125 to 4000 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.

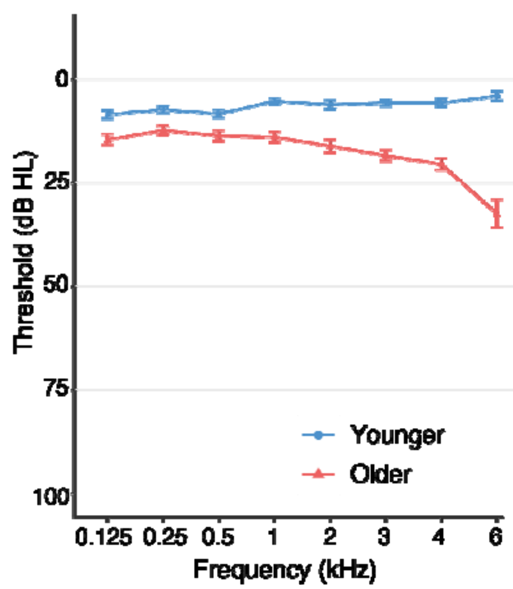

Figure 1. Audiogram of the grand average across ears for younger (blue) and older (red) participants. Error bars indicate ± one standard error. Both groups have clinically normal hearing, defined as pure-tone thresholds ≤ 25 dB HL from 125 - 4000 Hz in at least one ear, and no more than 10 dB difference between ears at each frequency.

Stimuli and experimental design

One-minute-long audio stimulus segments (22.05 kHz sampling rate) were constructed with MATLAB (MathWorks) as described in (Ding and Simon 2012b) and were all presented diotically (identically in each ear). 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. Four types of stimuli were presented: single talker ("quiet speech"), two talkers ("competing talkers") at two different relative loudness levels (0 dB SNR and -6 dB SNR), and single talker mixed with three-talker babble ("babble speech"). For the competing talker speech trials, participants were asked to selectively attend to one talker while ignoring the other, for which there were two signal-to-noise ratio (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 level. 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 talkers 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 babble speech condition, only the attended female talker speech was presented); otherwise, no speech segment was ever re-used across blocks.

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

To motivate the participants 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, participants 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

Non-invasive neuromagnetic responses were recorded using 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 anti-aliasing low-pass filter with cut off frequency at 500 Hz and a 60 Hz notch filter. Three separate channels function 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 was recorded simultaneously with and 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 was first denoised by removing non-functioning and saturated channels, and then with Time Shifted Principle Component Analysis (TSPCA) (de Cheveigné and Simon 2007), using the three reference channels to project out environmental magnetic noise not related to the brain activity, and then with Sensor Noise Suppression (SNS) (de Cheveigné and Simon 2008a) to project out sensor specific noise. To focus on low frequency cortical activity, the remaining data was band pass filtered between 1-10 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) (Sarela and Valpola 2005; de Cheveigné and Simon 2008) 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 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. 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 Biesmans et al. (2017). Then the envelope was filtered with the same bandpass filter (1–10 Hz) applied to the MEG data (and group delay compensation) and downsampled to 250 Hz and 100 Hz for TRF or 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 Institute of Health Toolbox (Gershon et al. 2013) was used as a measure of the subject's general behavioral ability to suppress competing stimuli in a visual scene. Participants 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 (congruent) or different (incongruent) to 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 (SPIN) 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. Due to the COVID-19 pandemic, only data from 32 subjects were obtained: 18 younger adults and 14 older adults. The task was run via from a graphical user interface in MATLAB. Subjects listened to 3–5 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). Participants were asked to repeat back the speech segment, in the case of quiet speech, and the selectively attended speech segmented, otherwise. At each noise level there were six different speech segments; the first trial 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, in the same order presented as in the MEG study for that subject.

QuickSIN Test. The Quick Speech-in-Noise test (QuickSIN) (Killion et al. 2004), a standardized measure of a listener's ability to understand speech in noise, was also employed. Due to 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 $E_{att}(t)$ and $E_{unatt}(t)$ respectively. For each subject and each trial, the attended and unattended speech envelopes were reconstructed separately (but simultaneously) using a linear temporal decoder applied to the first 6 DSS components ($D(d,t)$) estimated by the Boosting algorithm (David et al. 2007; Ding and Simon 2013; Ding et al. 2014) as follows.

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

Where $\varepsilon(t)$ 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). 10-fold 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 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 (TRF)

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 (Haufe et al. 2014). 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 (Lalor et al. 2009; Ding and Simon 2012b), 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 $E_{att}(t)$ and $E_{unatt}(t)$ respectively, with the model is formulated as:

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

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. $h_{att}(\tau)$ and $h_{unatt}(\tau)$, 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 using the boosting algorithm with 10-fold cross-validation, to minimize the mean square error between the predicted neural response and the true neural response (David et al. 2007; Zion Golumbic et al. 2013). 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 as the $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ respectively. Each peak corresponds to a different stage in the auditory signal processing chain. Latency and polarity of these peaks can be compared to the P1, N1 and P2 of standard cortical auditory evoked potentials (CAEP). For each subject and condition, peak latencies were extracted within a specific time range; for $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ 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 (R Core Team 2020). Linear mixed effect models (LMM) were used to systematically evaluate the relationships between the dependent (behavioral scores, neural measures) and independent variables (age, noise level, selective attention). For the LMM analysis, the *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2017) packages in R were used. The best fit model from the initial full model was found by the *buildmer* package (Voeten 2020) 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 backwards elimination stepwise procedure to identify the model of best fit for random and fixed effects. The assumptions of mixed effect modelling, 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 using Satterthwaite approximation for degrees of freedom (Satterthwaite 1941; Luke 2017). In order to interpret significant fixed effect interaction terms, variables were releveled to obtain model estimates for individual factor levels (indicated by 'with [*new reference level*] reference level'). The summary tables for each model used in result section are reported in the *Supplementary Materials*.

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 level* 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 level* could vary across subjects (random intercept for *subject* and random slopes for *noise level* by *subject* respectively). The full model for each dependent variable was defined as *intelligibility* ~ *age* × *noise level* + (*noise level* |*subject*). To evaluate the relationship between the two measures, data were analyzed using a separate LMM with *SPIN score* as the dependent variable, intelligibility rating as the independent variable and *subject* as a random intercept: SPIN score ~ *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 level*, 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 level on the neural features measured for the attended talker, 2) to examine the effects of aging, noise level and attention on the neural features measured for the attended talker and unattended talker. Only data from two competing talkers noise conditions were used for the latter model. The full models for 1) and 2) were defined as*: neural feature* ~ *age* × *noise level* + (*1* + *noise level* |*subject*) and *neural feature* ~ *age* × *noise level* × *attention* + (*1* + *noise level* × *attention* |*subject*), respectively.

To model nonlinear changes in the integration window analysis, Generalized Additive Mixed Models (GAMMs) (Wood 2006) in R (packages *mgcv*, *itsadug*) were used to analyze the reconstruction accuracies over integration window. Compared to Generalized Linear Mixed Effect Models, GAMMs have several advantages for modeling time series data, especially in electrophysiology (DeCat et al. 2014; Tremblay and Newman 2015) and pupillometry (van Rij et al. 2019). In particular, it 1) can model both linear and non-linear 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 to conventional non-linear mixed effect modelling where non-linear 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 parametric term (age, noise level and attention) and corresponding smooth term, model testing was compared between a test model and baseline model using Chi-Square (*compareML* in the *itsadug* package) to determine the significance of predictors (van Rij et al. 2019; Sóskuthy 2017). The parametric term indicates the overall difference in height between two curves, whereas smooth terms indicate the difference in shape (i.e., "wiggliness") 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 modelling and model diagnostics (over-smooth or under-smooth) were performed by residual plots and the *gam.check()*function.

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 for 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 level on the objective intelligibility measures (SPIN scores) were analyzed using LMM. Figure 2(a) plots the SPIN scores for both groups at all noise levels. The best fit model included main effects of noise level on speech SPIN scores and random intercept by subject (1). There was no effect of age on the SPIN scores. SPIN scores significantly dropped with increasing noise level (Quiet to 0 dB to -6 dB to Babble) with the highest drop from 0 dB to -6 dB (with Quiet reference level: *noise level*(0 dB): β = -12.88, *SE* = 1.57, p < 0.001; with 0 dB reference level: *noise level*(-6 dB): β = -40.95, *SE* = 1.57, p < 0.001; with -6 dB reference level: *noise level*(Babble): β = -13.52, *SE* = 1.92, p < 0.001). The lack of significant age effects will be addressed in discussion.

Intelligibility ratings. Parallel LMM analysis of the subjective intelligibility ratings Figure 2(b) revealed fixed effects of both *age* and *noise level* along with random slopes for *noise level* by *subject* (1). Older adults rated the intelligibility slightly higher compared to younger adults (β = 0.76, *SE* = 0.36, t = 2.11, p = 0.04). As in the case of SPIN scores, intelligibility ratings dropped significantly with worsening noise level in both groups (with Quiet reference level: *noise level*(0 dB): β = -2.6, *SE* = 0.27, p < 0.001, with 0 dB reference level: *noise level*(-6 dB): β = -1.51, *SE*

= 0.27, p < 0.001, with -6 dB reference level: *noise level*(Babble): = -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 (= 8.2, *SE* = 0.58, *p* < 0.001) and revealed no effects of age.

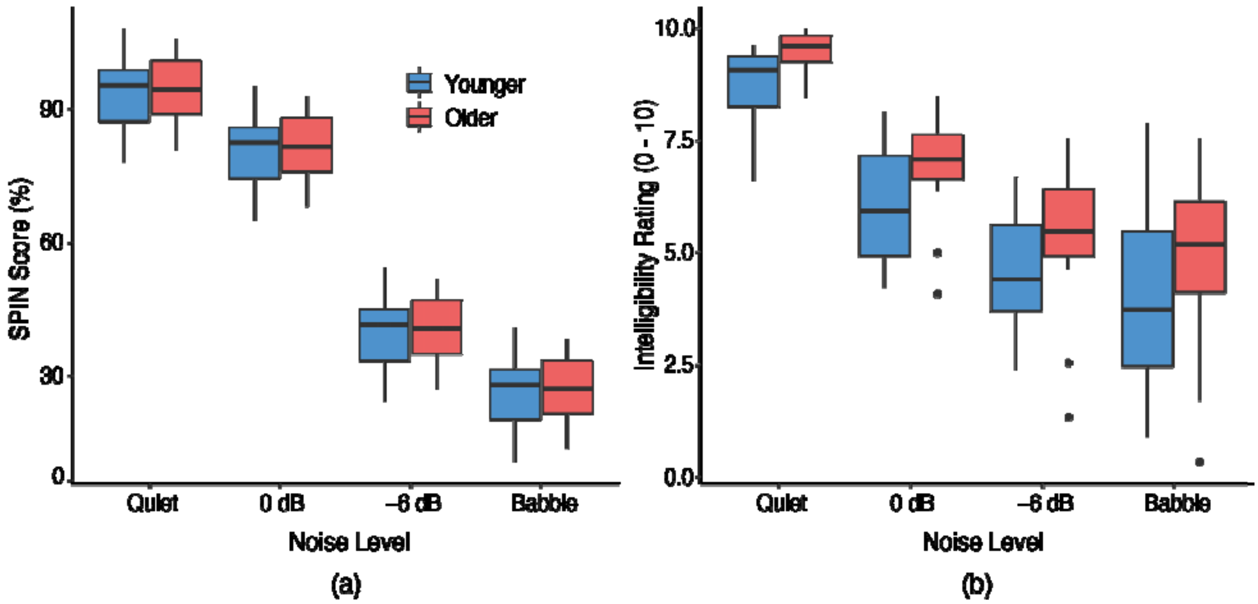

Figure 2. Behavioral test results for (a) speech SPIN scores (0–100%) and (b) intelligibility ratings (0–10). Both scores drop significantly with the noise level. A significant effect of age was seen only for the intelligibility rating whereas no age effect was found on the SPIN scores

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 level on the attended talker envelope tracking. As summarized in Table S3 the best fit model revealed main effects of *age*, *noise level* and *age* × *noise level* interactions with random intercepts by subject. For both groups and for all noise levels, the reconstruction accuracies fitted by the model for the attended talker are shown in Figure 3a. The main effects of *age* revealed that aging is associated with higher reconstruction accuracies (= 0.04, *SE* = 0.01, *p* < 0.001) in all noise levels.

The significant interaction *age* × *noise level* term revealed that the aging adversely affects speech reconstruction from quiet to noisy speech (*age*(Older)×*noise level*(0 dB): β = -0.018, *SE* = 0.01, *p* = 0.01, *age*(Older)×*noise level*(-6 dB): β = -0.015, *SE* = 0.01, *p* = 0.03, *age*(Older)×*noise level*(Babble): β = -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)×*noise level*(Babble): β = -0.005, *SE* = 0.01, *p* = 0.58). As can be seen from Figure 3a, the main effect of noise level revealed that the attended talker envelope reconstruction accuracies significantly reduce from quiet to noisy conditions in both groups (*noise level*(0 dB): β = -0.02, *SE* = 0.004, p < 0.001, *noise level*(-6 dB): β = -0.02, *SE* = 0.004, *p* < 0.001, *noise level*(Babble): β = -0.05, *SE* = 0.005, *p* < 0.001). No significant difference was observed between the 0 dB and -6 dB noise levels (with 0 dB reference level: *noise level*(-6 dB): β = 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 level*(Babble): β = -0.03, *SE* = 0.005, *p* < 0.001, with -6 dB reference level: *noise level*(Babble): β = -0.03, *SE* = 0.005, *p* < 0.001).

Secondly, to investigate the combined effects of selective attention, age, and noise level, 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 Figure 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* (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 (β = 0.03, *SE* = 0.01, *p* < 0.001). Furthermore, in both groups the attended talker was better represented than the unattended talker (β = -0.03, *SE* = 0.004, *p* < 0.001). For attended talker vs. unattended talker, or for either of the age groups, no significant difference was observed between the 0 dB and -6 dB noise levels.

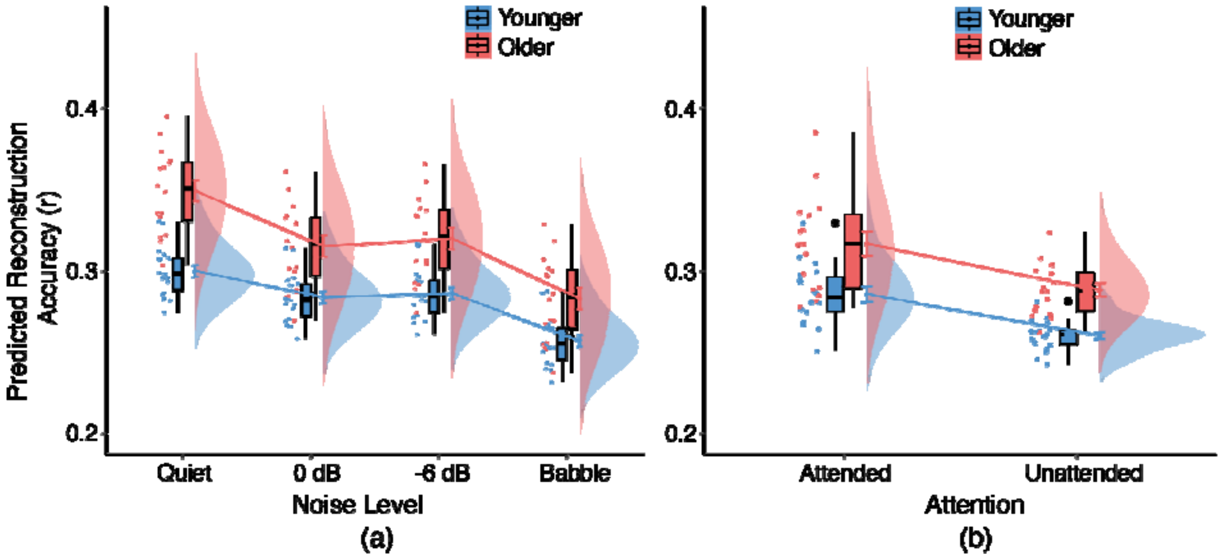

Figure 3. Model-predicted values for reconstruction accuracy for (a) the attended speech for both younger and older adults and for all noise levels, and (b) the attended vs unattended speech envelope reconstruction accuracies for competing talker conditions (there is no separation by noise level since no significant dependence on noise level was found). Both attended and unattended speech envelope reconstruction illustrates 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 to unattended.

Integration Window Analysis

Integration window analysis was done using 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* × *noise level* × *attention* significantly improve the model fit ((23) = 212.1, $p < 0.001$). Moreover, adding the parametric fixed effect term, characterizing the height of these functions, *age* × *noise level* × *attention*, and random smooth per subjects, also improved the model fit ((32) = 1531.39, $p < 0.001$,   (34) = 1886.42, $p < 0.001$ 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 the supplementary document (Table S10). Figure 4(a) 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 4(b) shows the dynamical differences between attended and unattended talker reconstruction accuracy curves for both younger and older adults and for the 0 dB noise level. The color-coded horizontal lines at graph bottom indicate where the differences are significant. In both age groups, attended talker reconstruction accuracies were significantly higher compared to unattended talker after the middle processing stage (~150 ms). Interestingly, in older adults, for the -6 dB noise level, the unattended talker representation is enhanced compared to 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 (Figure 4(c)) and 55 ms for the unattended talker when averaged across noise levels. As can be seen from Figure 4(c), 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 to younger adults until later processing stages.

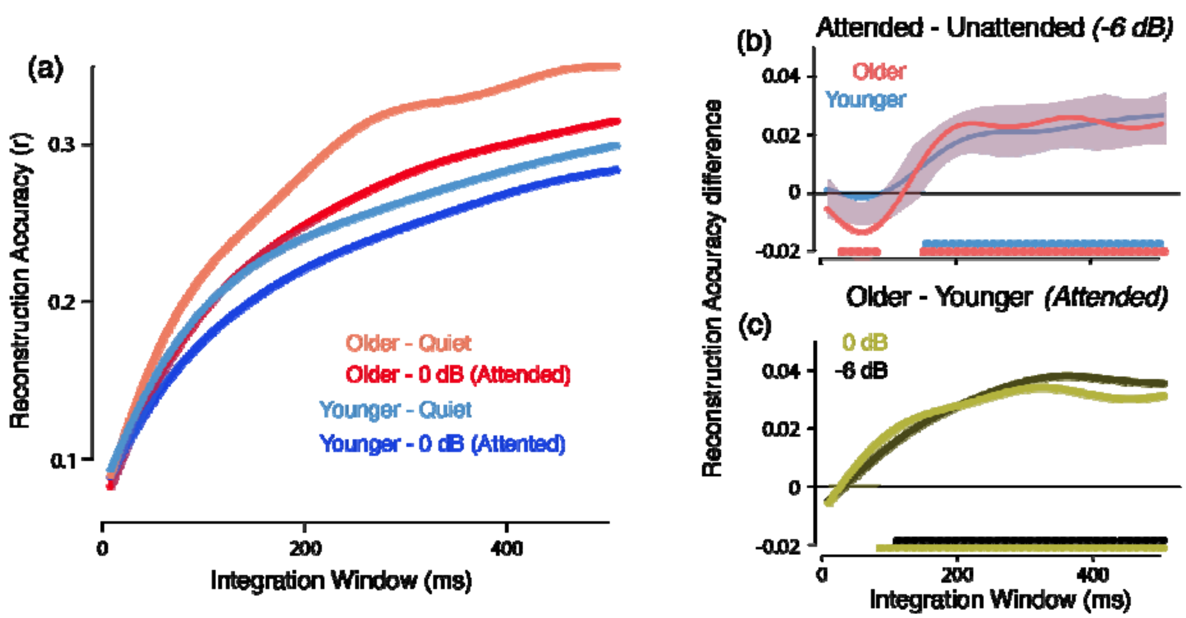

Figure 4. Integration Window Analysis using GAMMs. (a) Changing reconstruction accuracy with integration window (only a subset of curves is shown, for visual clarity). (b) Reconstruction accuracy difference between attended vs unattended talker for -6 dB noise level. (c) Reconstruction accuracy difference between older vs younger for attended talker. The color-coded horizontal lines above the horizontal axis in (b, c) mark where the difference is statistically significant. 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 attended talker envelope starts at early processing stages (<100 ms).

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 ($M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$) we separately analyzed individuals' peak amplitude and latency, comparing between the two age groups, and for both attended and unattended talker and noise levels. Figure 5(a) visualizes how the TRF peak amplitudes and latencies varied for the attended talker across two age groups and noise levels. Figure 6(a) visualizes how the TRF peak amplitudes and latencies varied between attended vs unattended talker.

TRF peak amplitudes

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

TRF peak amplitudes for the $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ are plotted in Figure 5(a, b). Overall, in the comparison between younger vs older, older adults showed exaggerated peak amplitudes in all 3 processing stages ($M50_{TRF}$: *age*(Older): β = -0.01, *SE* = 0.004, *p* = 0.03, $M100_{TRF}$: *age*(Older): β = -0.02, *SE* = 0.005, *p* = 0.008 and $M200_{TRF}$: *age*(Older): β = -0.018, *SE* = 0.003, *p* = 0.001). For both the $M50_{TRF}$ and $M200_{TRF}$, peak amplitudes were stronger in all noise levels and that was significant only for the quiet speech condition ($M50_{TRF}$: *age*(Older): β

= 0.02, *SE* = 0.005, $p < 0.001$, $M200_{TRF}$: *age*(Older): $\beta = 0.03$, *SE* = 0.01, $p < 0.001$). In contradistinction, except for the quiet speech, the $M100_{TRF}$ was stronger in all the noisy conditions (*age*(Older): $\beta = 0.01$, *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 level revealed that the $M50_{TRF}$ response decreases with the task difficulty in both groups (*noise level*(0 dB): $\beta = -0.01$, *SE* = 0.003, $p = 0.01$, *noise level*(-6 dB): $\beta = -0.01$, *SE* = 0.003, $p = 0.04$, *noise level*(Babble): $\beta = -0.02$, *SE* = 0.003, $p < 0.001$, with Older reference level: *noise level*(0 dB): $\beta = -0.01$, *SE* = 0.003, $p < 0.001$, *noise level*(-6 dB): $\beta = -0.2$, *SE* = 0.003, $p < 0.001$, *noise level*(Babble): $\beta = -0.03$, *SE* = 0.003, $p < 0.001$). A significant *age* × *noise level* interaction indicated that aging contributes more to the $M50_{TRF}$ reduction as a function of noise level (*age*(Older) ×*noise level*(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 level the $M100_{TRF}$ significantly increased (with Older reference level: *noise level*(0 dB): $\beta = 0.01$, *SE* = 0.003, $p < 0.001$), while the $M200_{TRF}$ decreased (with Older reference level: *noise level*(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 level*(Babble): $\beta = -0.01$, *SE* = 0.003, $p = 0.02$, with Older, -6 dB reference level: *noise level*(Babble): $\beta = -0.01$, *SE* = 0.003, $p = 0.003$).

To investigate the combined effects of aging, selective attention and noise level on the TRF amplitude responses, separate LMM models were constructed (Table S6). Figure 6(b) 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 to 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*, *attention* and an *age* × *attention* interaction effect. Compared to 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): β = 0.01, *SE* = 0.01, p = 0.19). In both age groups unattended peak amplitudes were reduced compared to attended peak amplitudes (*attention*(Unattended): β = -0.01, *SE* = 0.003, $p < 0.001$, with Older reference level: *attention*(Unattended): β = -0.02, *SE* = 0.003, $p < 0.001$), and the interaction effect revealed that this reduction in peak heights is amplified by aging (β = -0.01, *SE* = 0.004, $p < 0.001$). Interestingly, the $M200_{TRF}$ showed main effects of *age* and *attention*, and an *attention* × *noise level* interaction. $M200_{TRF}$ amplitudes in both attended and unattended TRFs were stronger in older adults (*age*(Older): β = 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 to the unattended (*attention*(Unattended): β = -0.01, *SE* = 0.002, $p < 0.001$).

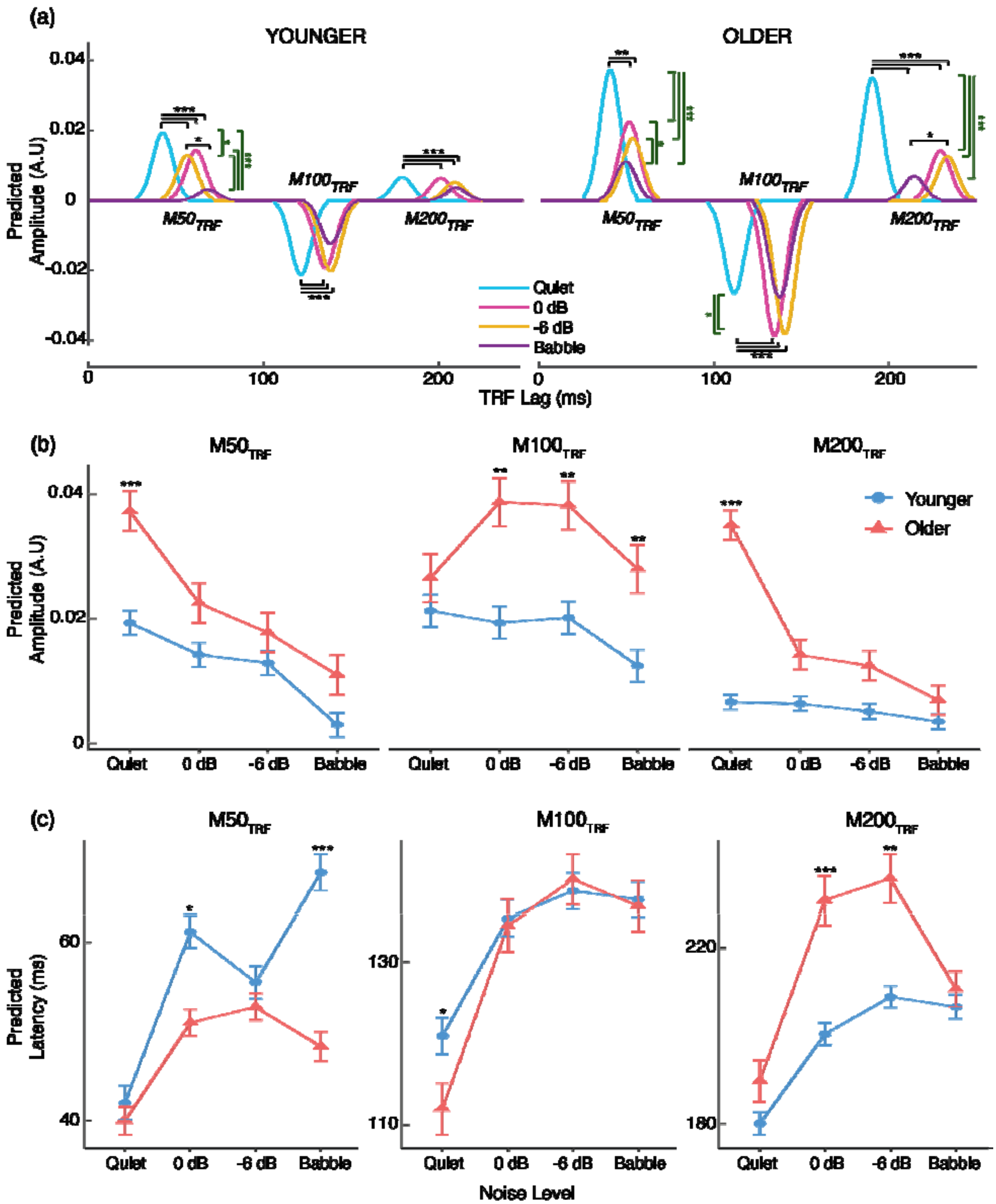

Figure 5. Model-predicted attended talker TRF peak amplitudes and latencies. (a) TRFs showing overall amplitudes and latency for both groups and all noise levels (for visualization simplicity, all peaks are represented with the same Gaussian shape, standard deviation 7 ms, centered at

*the group mean peak latency, and with amplitude given by the group mean peak amplitude). (b) The TRF amplitudes (±SE) as a function of noise level. Generally, older adults (red) exhibit stronger TRF peak amplitudes compared to 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, the $M100_{TRF}$ increases and $M200_{TRF}$ decreases but only in older adults. (c) The TRF latencies (±SE) as a function of noise level. Compared to 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. *$p<0.05$, **$p<0.01$, ***$p<0.001$*

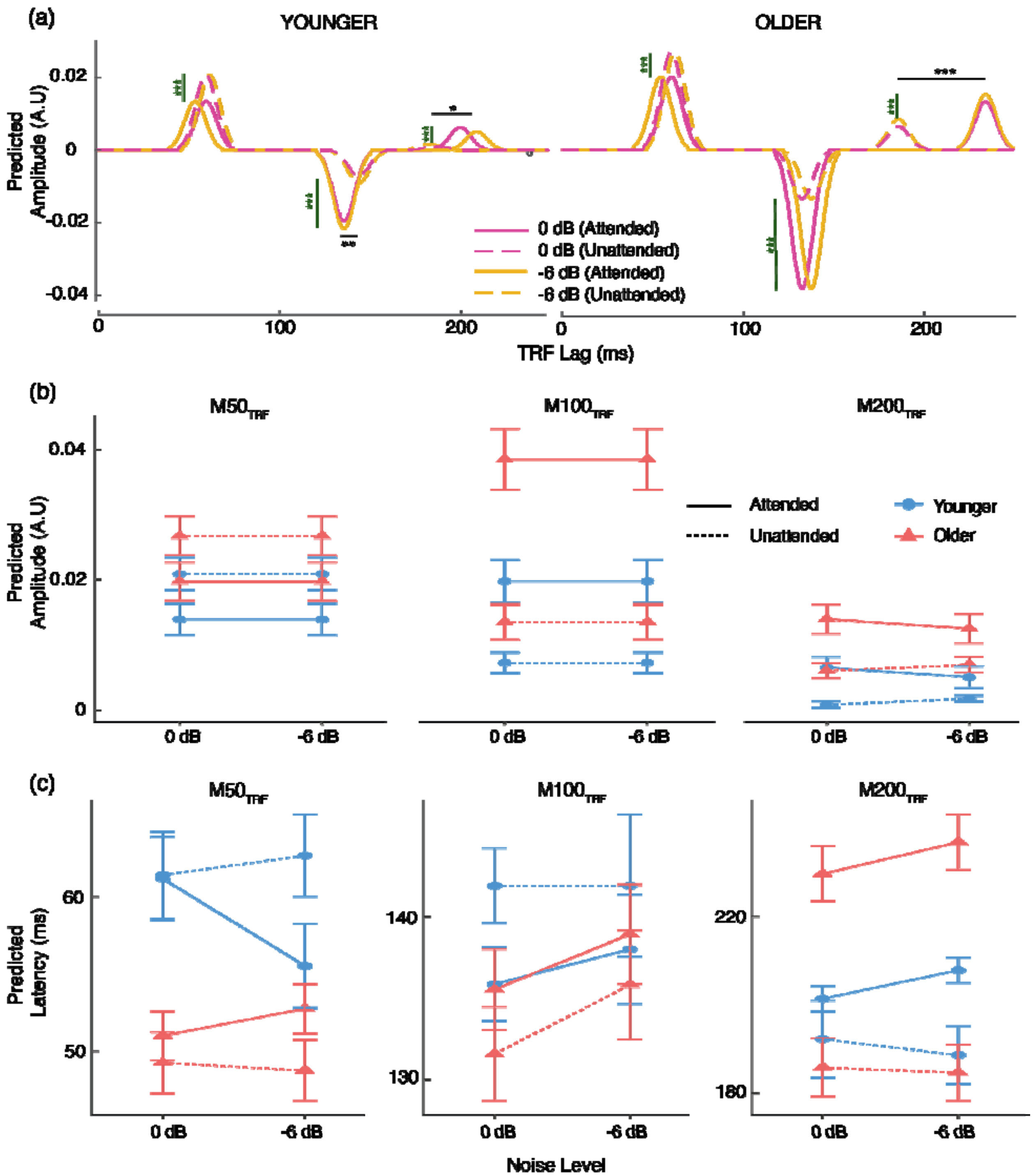

Figure 6. Model-predicted values for attended vs unattended talker TRF peak amplitude and latency. (a) TRF peak amplitudes and latency for both groups and two talker speech conditions for both attended and unattended talker (solid line = Attended, dashed line = Unattended). For visualization simplification, peaks are represented with a common Gaussian shape as in Figure 5. (b) TRF amplitudes (±SE) as a function of noise level. Generally older adults exhibit stronger

*TRF amplitudes for both attended and unattended peaks. Attended $M50_{TRF}$ is significantly smaller compared to unattended amplitudes. In contrast, attended $M100_{TRF}$ and $M200_{TRF}$ are enhanced compared to unattended peak amplitude. (c) TRF latencies (±SE) as a function of noise level. The attended $M200_{TRF}$ peak is significantly delayed compared to the unattended $M200_{TRF}$ peak and this difference is bigger in older adults. *p<0.05, **p<0.01, ***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 level* and *age* × *noise level* along with random intercepts by subject for $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ (Table S7). Model predicted latencies are plotted in Figure 5(c). Averaged latencies over noise levels revealed that compared to 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 level*(0 dB): $\beta = 18.48$, $SE = 3.02$, $p < 0.001$, $M100_{TRF}$: *noise level*(0 dB): $\beta = 14.59$, $SE = 2.75$, $p < 0.001$, $M200_{TRF}$: *noise level*(0 dB): $\beta = 21.80$, $SE = 4.64$, $p < 0.001$) and older adults (with Older reference level: $M50_{TRF}$: *noise level*(0 dB): $\beta = 10.83$, $SE = 3.09$, $p < 0.001$, $M100_{TRF}$: *noise level*(0 dB): $\beta = 22.13$, $SE = 2.92$, $p < 0.001$, $M200_{TRF}$: *noise level*(0 dB): $\beta = 38.80$, $SE = 4.86$, $p < 0.001$) highlighting that the peak responses are delayed with the stimulus noise level. With respect to quiet, babble speech latencies in all three processing were delayed in both younger (*noise level*(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 level*(Babble): $M50_{TRF}$: $\beta = 9.24$, $SE = 3.23$, $p = 0.006$, $M100_{TRF}$: $\beta = 25.63$, $SE = 2.87$, $p < 0.001$, $M200_{TRF}$: $\beta = 24.59$, $SE = 4.98$, $p < 0.001$). The *age* × *noise* level interaction manifests as aging contributing more to the peak delay from quiet to 0 dB for both $M100_{TRF}$ (*age*(Older) × *noise level*(0 dB): $\beta = 8.05$, $SE = 3.98$, $p = 0.04$) and $M200_{TRF}$ (*age*(Older) × *noise level*(0 dB): $\beta = 17.08$, $SE = 6.68$, $p = 0.01$).

The effect of selective attention on TRF peak latencies was analyzed using LMMs (Table S8) and results are displayed in Figure 6(c). Comparing mean latencies across noise levels revealed that compared to 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 to 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 to the unattended peak (*attention*(Unattended): $\beta = -12.02$, $SE = 6.28$, $p = 0.05$), and this effect was stronger in older adults (*age*(Older)×*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 $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ separately. Peak amplitudes were predicted by age and peak latencies and random intercepts by subject. As can be seen from Figure 7, older adults' $M200_{TRF}$ peak amplitudes exhibited a significant negative relationship with the latency (with Older reference level: *Latency*: $\beta = -0.0003$, $SE = 0.0001$, $p < 0.001$), 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*: $\beta = -0.0001$, $SE = 0.0001$, $p = 0.36$) (Table S9).

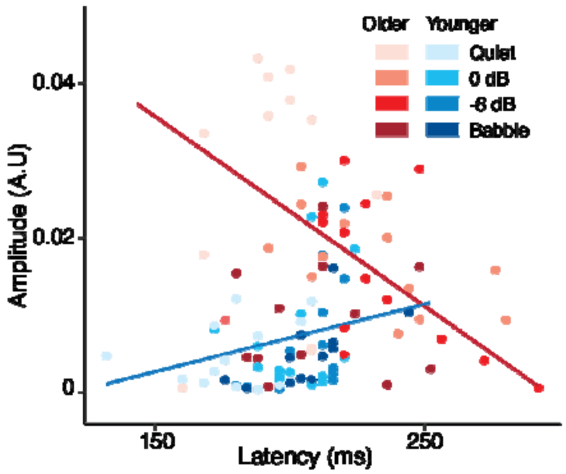

Figure 7. M200$_{TRF}$ peak amplitude vs latency. Four different shades of each color represent each noise levels. Older adults' amplitudes were negatively associated with their latencies.

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. Firstly, we analyzed how attended talker neural features are affected by speech intelligibility score and age but without specific regard to stimulus noise level. Results revealed that the reconstruction accuracy increases with better speech intelligibility in both groups (*SPIN score*: = 0.0005, *SE* = 0.001, $p < 0.001$) (Figure 8(a)). Similarly, TRF peak amplitude analysis revealed that stronger M50$_{TRF}$ amplitudes are associated with better speech intelligibility score (*SPIN score*: = 0.0002, *SE* = 0.001, $p < 0.001$) (Figure 8(b)), and this effect was stronger in older adults (*age*(Older)×*SPIN score*: = 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*: = 0.0003, *SE* = 0.001, $p < 0.001$), whereas no significant trend was found for younger adults (*SPIN score*: = 0.0001, *SE* = 0.001, $p = 0.022$) (Figure 8(c)). 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$_{TRF}$: = -0.17, *SE* = 0.04, $p < 0.001$, M100$_{TRF}$: = -0.22, *SE* = 0.03, $p < 0.001$, M200$_{TRF}$: = -0.32, *SE* = 0.06, $p < 0.001$). Additionally, similar trends were observed when subjective intelligibility rating was used instead the SPIN scores.

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

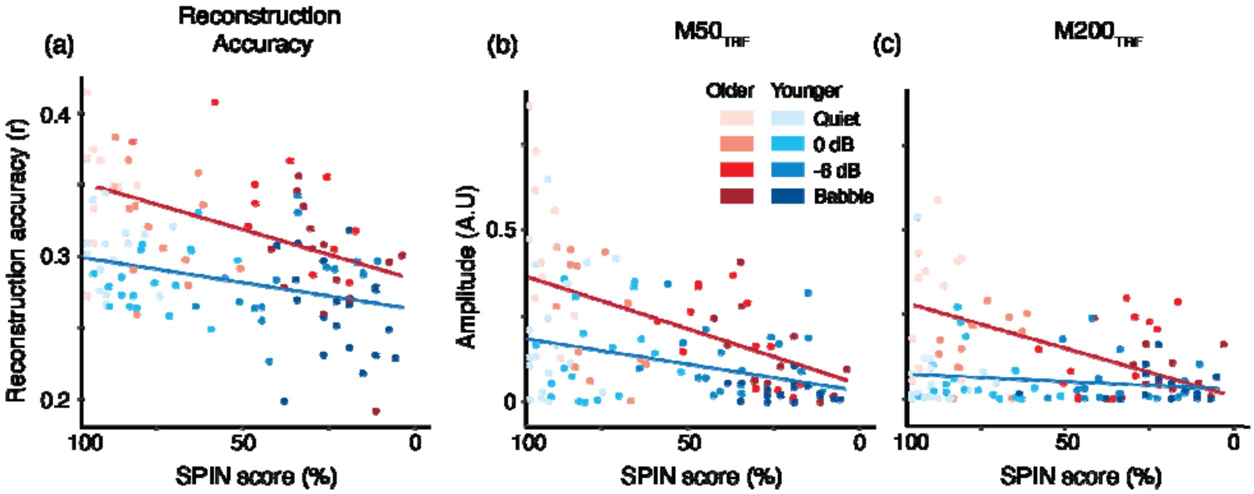

Figure 8. Neural measure vs SPIN scores including all noise levels. (a) Reconstruction accuracy vs SPIN score. (b) M50$_{TRF}$ peak amplitude vs SPIN score. Better reconstruction accuracies or M50$_{TRF}$ peak amplitudes were related to better speech intelligibility scores in both groups. (c). M200$_{TRF}$ peak amplitude vs SPIN score. Only in older adults, stronger M200$_{TRF}$ peak amplitude was associated with better speech intelligibility score.

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 (Decruy et al. 2019; Presacco et al. 2016a) and further investigated the cortical processing stages associated with this exaggeration. Using the integration window analysis and TRF peaks, we showed 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 also show that aging enhances this reduction. In particular, TRF peak analysis 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 revealed that all processing stages were delayed with interfering speech, which was further affected by aging.

Aging is associated with exaggerated speech envelope representation/encoding

Perhaps counterintuitively, the reconstruction analysis demonstrated that compared to younger adults, older adults exhibit exaggerated speech envelope representation irrespective of the noise level. This replicates previous results by Presacco et al. (2016b) 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 nonspeech (Bidelman et al. 2014; Goossens et al. 2016; Irsik et al. 2021) and continuous speech stimuli (Decruy et al. 2018; Mesik et al. 2021). Whereas both Presacco et al. (2016b) and Decruy et al. (2018) analyzed the attended speech envelope reconstruction, the current 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 (Griffiths and Warren 2004), separable via neural implementations of auditory scene analysis (Shinn-Cunningham 2008; Shamma et al. 2011). Here we 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, $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$, and for both attended and unattended speech. The enhanced $M50_{TRF}$ 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 using both EEG (McCullagh and Shinn 2013; Roque et al. 2019b) and MEG (Brodbeck et al. 2018; Zan et al. 2020), for both speech in quiet and in noisy conditions. Alain et al. (2014) suggested that this increased neural activity may be caused by excitatory and inhibitory imbalance, which is further in agreement with animal studies (McCullagh and Shinn 2013), and is consistent with other studies (Brodbeck et al. 2018). Larger cortical peaks at ~100 ms latency (e.g., the $M100_{TRF}$) in older adults have been reported using both EEG (McCullagh and Shinn 2013) and MEG (Zan et al. 2020), and the exaggeration is bigger for the attended speech compared to unattended. The exaggerated response at this middle processing stage has been associated with increased task-related effort (Rao et al. 2010). For longer latency cortical peaks (~200 ms latency, e.g., the $M200_{TRF}$), previous studies have reported an enhanced late peak in both EEG (O'Sullivan et al. 2015; Fiedler et al. 2019) and MEG (Zan et al. 2020) when the stimulus was continuous speech. No age-related enhancement was seen for this stage, however, for a gap-in-noise detection task (Alain and Snyder 2008; Lister et al. 2011). 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, due to the importance of inhibition for neuronal tuning, likely worse sensitivity, both temporally and spectrally. Gamma-aminobutyric acid (GABA)-mediated inhibition plays a major role in maintaining synchrony and spectral sensitivity in auditory circuits. Both animal (Hughes et al. 2010; Caspary et al. 2013; Richardson et al. 2013; Parthasarathy et al. 2019; Ramamurthy and Recanzone 2020) and human (Lalwani et al. 2019; Dobri and Ross 2021) 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. (2010) 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 (Wild et al. 2012; Pichora-Fuller et al. 2016; Rumschlag et al. 2022). 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 Kuchinsky and Vaden (2020). 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., de-differentiation, predictive of poor performance; for a review see Wingfield and Grossman (2006))

Attentional modulation of speech representation/encoding and aging

In line with the results for younger adults (Mesgarani and Chang 2012; Ding and Simon 2012; O'Sullivan et al. 2015; Das, Bertrand, and Francart 2018), we found that in older adults the attended speech envelope is better represented than the unattended. Surprisingly, irrespective of exaggerated envelope representation in the cortex, both groups showed similar effects of selective attention on the envelope representation as no age by selective attention interaction effect was found.

Both integration window analysis and TRF analysis contrasting attended vs unattended speech reveal that the unattended speech is represented to a similar degree as the attended (or even more strongly in the case of -6dB 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 (though see Brodbeck et al. 2020). Older adults exhibit an enhanced attended-unattended $M100_{TRF}$ amplitude difference compared to 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 the previous studies, which report that envelope tracking is adversely affected by the SNR (Ding and Simon 2013; Presacco et al. 2016a,b; Das et al. 2018). Our results also show that older adults' envelope representation is more strongly affected, still negatively, by adding higher levels of noise, compared to younger adults. Thus, within an individual, better reconstruction accuracy is associated with clearer speech, but across subjects, better reconstruction accuracy is not necessarily associated with better speech intelligibility, and especially when comparing across age groups, where the association may change sign.

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 (Fiedler et al. 2019; Brodbeck et al. 2020). The older adults showed greater reduction in $M50_{TRF}$ amplitude with the noise level, 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 level, whereas no significant $M100_{TRF}$ amplitude differences are seen between noise levels in younger adults. These results are consistent with some previous studies (Ding and Simon 2012b; McCullagh et al. 2012; Rufener et al. 2014), but not all (McCullagh and Shinn 2013; Billings et al. 2015; Zan et al. 2020) where a dependence on masker noise level is found in both groups. Mechanistically, an exaggerated $M100_{TRF}$ amplitude may reflect an increase in task-related attention or cognitive effort (Rao et al. 2010; Billings et al. 2015), and as such any conflicting trends maybe 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 (Pichora-Fuller 2008). 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* (Zan et al. 2020), though the current 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 contributes 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 (Ding and Simon 2013; O'Sullivan et al. 2015). Specifically, Power et al. (2012) and O'Sullivan et al. (2015) 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 Figure 4(d), as long as 350 ms (Presacco et al. 2016a). This provides support for an 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 (Tremblay et al. 2004; O'Brien et al. 2015), 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 (McCullagh and Shinn 2013). The latency results for babble speech were less clear. Latencies for the babble condition were delayed compared to quiet speech, yet no consistent trends were observed compared to 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 (Kuchinsky and Vaden 2020).

Compared to younger adults, older adults demonstrated relatively early $M50_{TRF}$ peaks. This effect has been observed in some studies using both CAEP (O'Brien et al. 2015; Roque et al. 2019b) and MEG (Brodbeck et al. 2018), but not others (Tremblay et al. 2004; Alain et al. 2014). This finding is consistent with an excitation/inhibition imbalance favoring excitation compared to younger adults. Another possible explanation is that an $M50_{TRF}$ followed immediately by an exaggerated $M100_{TRF}$ of the opposite polarity would appear shortened due to an earlier cut off 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 (Tremblay and Newman 2015; Roque et al. 2019a), suggesting that 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 level in older adults suggesting over-reliance on the middle and late processing mechanisms in older adults to compensate for degraded afferent input (Parthasarathy et al. 2020). 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. (2019), 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 current 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 level increased in both groups, suggesting that noise level 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 which 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 (Pichora□Fuller et al. 1995; Pichora-Fuller 2008; Schneider et al. 2016) 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 (Killion et al. 2004) are known to show such behavioral age-related auditory declines (Presacco et al. 2016a,b; Holder et al. 2018). Unfortunately, due to the COVID-19 pandemic, QuickSIN measures were obtained from only half of the subjects, which were 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 the older adults tend to underestimate their hearing difficulties in comparison to younger adults (Uchida et al. 2003; Gosselin and Gagné 2011). It is unclear whether subjective judgments in older adults were influenced by intelligibility, contextual factors, or other variables.

The positive association between the behavioral SPIN scores and intelligibility ratings do suggest that models using (subjective) intelligibility ratings collected during the task of interest would give similar outcomes as those using (objective) SPIN scores collected in a separate task (Hazan et al. 2018). However, the age effects revealed only in subjective intelligibility ratings suggests 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 (reconstruction accuracy $M50_{TRF,}$ and TRF peak amplitudes) are related to the speech intelligibility which is similar to the observations in noise level effect. This is expected as the noise level and speech intelligibility are correlated. Unexpectedly, we did not find any consistent relationships between neural measures and behavioral performance measures within a noise level. 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.

CONCLUSION

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 processing stages; early, middle and late processing. Moreover, the effects of speech intelligibility and attention on $M50_{TRF}$, $M100_{TRF}$ and $M200_{TRF}$ peak amplitudes and latencies reveal characteristics related to different auditory processing stages, and aging appears to differently affect the individual processing stages. Earlier and enhanced processing of early stages support the hypothesis of an excitatory and inhibitory imbalance, whereas delayed and enhanced processing of late stages support the hypotheses of increased attention and late compensatory mechanisms in older adults. Overall, these findings support the theory that 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.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Institute on Aging (P01-AG055365) and the; National Institutes of Deafness and Other Communication Disorders (R01-DC014085, R01-DC019394). We thank Katie Brow, Julie Cohen and Carol Gorham for their help in recruiting and testing subjects.

DISCLOSURES

The authors declare no competing financial interests. 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 U.S. Government.

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

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, 2014.

Alain C, **Snyder JS**. Age-related differences in auditory evoked responses during rapid perceptual learning. *Clin Neurophysiol* 119: 356–366, 2008.

Anderson S, **Karawani H**. Objective evidence of temporal processing deficits in older adults. *Hear Res* 397: 108053, 2020.

Anderson S, **Parbery-Clark A**, **White-Schwoch T**, **Kraus N**. Aging Affects Neural Precision of Speech Encoding. *J Neurosci* 32: 14156–14164, 2012.

Bates D, **Mächler M**, **Bolker B**, **Walker S**. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* 67, 2015.

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.

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.

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.

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.

Boettcher FA, **Mills JH**, **Norton BL**. Age-related changes in auditory evoked potentials of gerbils. I. Response amplitudes. *Hear Res* 71: 137–145, 1993.

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.

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.

Caspary DM, **Hughes LF**, **Ling LL**. Age-related GABAA receptor changes in rat auditory cortex. *Neurobiol Aging* 34: 1486–1496, 2013.

Caspary DM, **Milbrandt JC**, **Helfert RH**. Central auditory aging: GABA changes in the inferior colliculus. *Exp Gerontol* 30: 349–360, 1995.

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.

de Cheveigné A, **Simon JZ**. Denoising based on time-shift PCA. *J Neurosci Methods* 165: 297–305, 2007.

de Cheveigné A, **Simon JZ**. Sensor noise suppression. *J Neurosci Methods* 168: 195–202, 2008a.

de Cheveigné A, **Simon JZ**. Denoising based on spatial filtering. *J Neurosci Methods* 171: 331–339, 2008b.

Chisolm TH, **Willott JF**, **Lister JJ**. The aging auditory system: anatomic and physiologic changes and implications for rehabilitation. *Int J Audiol* 42: 3–10, 2003.

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.

David SV, **Mesgarani N**, **Shamma SA**. Estimating sparse spectro-temporal receptive fields with natural stimuli. *Netw Comput Neural Syst* 18: 191–212, 2007.

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.

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.

Decruy L, **Vanthornhout J**, **Francart T**. Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hear Res* 393: 107961, 2020.

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.

Ding N, **Simon JZ**. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol* 107: 78–89, 2012a.

Ding N, **Simon JZ**. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc Natl Acad Sci* 109: 11854–11859, 2012b.

Ding N, **Simon JZ**. Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *J Neurosci* 33: 5728–5735, 2013.

Dobri SGJ, **Ross B**. Total GABA level in human auditory cortex is associated with speech-in-noise understanding in older age. *NeuroImage* 225: 117474, 2021.

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.

Eckert MA, **Keren NI**, **Roberts DR**, **Calhoun VD**, **Harris KC**. Age-related changes in processing speed: unique contributions of cerebellar and prefrontal cortex. *Front Hum Neurosci* 4: 10, 2010.

Fabiani M. It was the best of times, it was the worst of times: A psychophysiologist's view of cognitive aging. *Psychophysiology* 49: 283–304, 2012.

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.

Fitzgibbons PJ, **Gordon□Salant S**. Age effects on duration discrimination with simple and complex stimuli. *J Acoust Soc Am* 98: 3140–3145, 1995.

Frisina DR, **Frisina RD**. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res* 106: 95–104, 1997.

Fuglsang SA, **Dau T**, **Hjortkjær J**. Noise-robust cortical tracking of attended speech in real-world acoustic scenes. *NeuroImage* 156: 435–444, 2017.

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.

Goossens T, **Vercammen C**, **Wouters J**, **Wieringen A van**. Aging Affects Neural Synchronization to Speech-Related Acoustic Modulations. *Front Aging Neurosci* 8, 2016.

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.

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.

Gosselin PA, **Gagné J-P**. Older Adults Expend More Listening Effort Than Young Adults Recognizing Speech in Noise. *J Speech Lang Hear Res* 54: 944–958, 2011.

Griffiths TD, **Warren JD**. What is an auditory object? *Nat Rev Neurosci* 5: 887–892, 2004.

Haufe S, **Meinecke F**, **Görgen K**, **Dähne S**, **Haynes J-D**, **Blankertz B**, **Bießmann F**. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87: 96–110, 2014.

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.

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.

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.

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.

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.

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.

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. Springer International Publishing, p. 231–256.

Kuznetsova A, **Brockhoff PB**, **Christensen RHB**. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw* 82, 2017.

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.

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.

Lister JJ, **Maxfield ND**, **Pitt GJ**, **Gonzalez VB**. Auditory evoked response to gaps in noise: Older adults. *Int J Audiol* 50: 211–225, 2011.

Luke SG. Evaluating significance in linear mixed-effects models in R. *Behav Res Methods* 49: 1494–1502, 2017.

Manan HA, **Yusoff AN**, **Franz EA**, **Mukari SZ-MS**. Effects of Aging and Background Babble Noise on Speech Perception Processing: An fMRI Study. *Neurophysiology* 49: 441–452, 2017.

McCullagh J, **Musiek FE**, **Shinn JB**. Auditory Cortical Processing in Noise in Normal-Hearing Young Adults. *Audiol Med* 10: 114–121, 2012.

McCullagh J, **Shinn JB**. Auditory cortical processing in noise in younger and older adults. *Hear Balance Commun* 11: 182–190, 2013.

McDowd JM, **Shaw RJ**. Attention and Aging: A Functional Perspective. In: *The Handbook of Aging and Cognition*. Lawrence Erlbaum Associates, 2000, p. 755.

Mesgarani N, **Chang EF**. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485: 233–236, 2012.

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.

Näätänen R, **Winkler I**. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull* 125: 826–859, 1999.

O'Brien JL, **Nikjeh DA**, **Lister JJ**. Interaction of Musicianship and Aging: A Comparison of Cortical Auditory Evoked Potentials. *Behav Neurol* 2015: 1–12, 2015.

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.

Overton JA, **Recanzone GH**. Effects of aging on the response of single neurons to amplitude-modulated noise in primary auditory cortex of rhesus macaque. *J Neurophysiol* 115: 2911–2923, 2016.

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.

Parthasarathy A, **Herrmann B**, **Bartlett EL**. Aging alters envelope representations of speech-like sounds in the inferior colliculus. *Neurobiol Aging* 73: 30–40, 2019.

Parthasarathy A, **Kujawa SG**. Synaptopathy in the Aging Cochlea: Characterizing Early-Neural Deficits in Auditory Temporal Envelope Processing. *J Neurosci* 38: 7108–7119, 2018.

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.

Phatak SA, **Brungart DS**, **Zion DJ**, **Grant KW**. Clinical Assessment of Functional Hearing Deficits: Speech-in-Noise Performance. *Ear Hear* 40: 426–436, 2019.

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.

Pichora-Fuller MK, **Kramer SE**, **Eckert MA**, **Edwards B**, **Hornsby BWY**, **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.

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.

Plack CJ, **Barker D**, **Prendergast G**. Perceptual Consequences of “Hidden” Hearing Loss. *Trends Hear* 18: 233121651455062, 2014.

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: A late locus of attention to natural speech. *Eur J Neurosci* 35: 1497–1503, 2012.

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, 2016a.

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, 2016b.

R Core Team. R: A Language and Environment for Statistical Computing [Online]. R Foundation for Statistical Computing.https://www.R-project.org/.

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.

Rao A, **Zhang Y**, **Miller S**. Selective listening of concurrent auditory stimuli: An event-related potential study. *Hear Res* 268: 123–132, 2010.

Richardson BD, **Ling LL**, **Uteshev VV**, **Caspary DM**. Reduced GABAA Receptor-Mediated Tonic Inhibition in Aged Rat Auditory Thalamus. *J Neurosci* 33: 1218–1227, 2013.

van Rij J, **Hendriks P**, **van Rijn H**, **Baayen RH**, **Wood SN**. Analyzing the Time Course of Pupillometric Data. *Trends Hear* 23: 233121651983248, 2019.

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, 2019a.

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, 2019b.

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.

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.

Sarela J, **Valpola H**. Denoising Source Seperation. *J Mach Learn Res* 233–272, 2005.

Satterthwaite FE. Synthesis of variance. *Psychometrika* 6: 309–316, 1941.

Schneider BA, **Avivi-Reich M**, **Leung C**, **Heinrich A**. How Age and Linguistic Competence Affect Memory for Heard Information. *Front Psychol* 7: 618, 2016.

Schneider BA, **Pichora-Fuller MK**. Implications of Perceptual Deterioration for Cognitive Aging Research. In: *The Handbook of Aging and Cognition*. Lawrence Erlbaum Associates, 2000, p. 755.

Shamma SA, **Elhilali M**, **Micheyl C**. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci* 34: 114–123, 2011.

Shinn-Cunningham BG. Object-based auditory and visual attention. *Trends Cogn Sci* 12: 182–186, 2008.

Snell KB. Age-related changes in temporal gap detection. *J Acoust Soc Am* 101: 2214–2220, 1997.

Sóskuthy M. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. *ArXiv Prepr ArXiv170305339 Stat* , 2017.

Strawbridge WJ, **Wallhagen MI**, **Shema SJ**, **Kaplan GA**. Negative Consequences of Hearing Impairment in Old Age. *The Gerontologist* 40: 320–326, 2000.

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.

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.

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 (Stockh)* 123: 618–626, 2003.

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.

Voeten CC. buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression [Online]. https://CRAN.R-project.org/package=buildmer.

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 Off J Soc Neurosci* 32: 14010–14021, 2012.

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.

Wingfield A, **Grossman M**. Language and the Aging Brain: Patterns of Neural Compensation Revealed by Functional Brain Imaging. *J Neurophysiol* 96: 2830–2839, 2006.

Wood SN. *Generalized additive models: an introduction with R*. Boca Raton, FL: Chapman & Hall/CRC, 2006.

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.

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.

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.

Supplementary Materials

Table S1. LMEM summary tables for behavioral analysis

	SPIN score				Intelligibility Rating			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***	***Estimates***	***SE***	***t value***	***p value***
Intercept [Younger, Quiet]	93.59	1.92	48.6	**<0.001**	8.72	0.27	32.2	**<0.001**
noise level [0 dB]	-12.88	1.57	-8.2	**<0.001**	-2.60	0.27	-9.5	**<0.001**
noise level [-6 dB]	-53.83	1.57	-34.3	**<0.001**	-4.10	0.34	-12.0	**<0.001**
noise level [Bab]	-67.35	1.92	-35.1	**<0.001**	-4.63	0.43	-10.9	**<0.001**
age [Old]					0.76	0.36	2.1	**0.035**
Random Effects	**Variance**	**SD**			**Variance**	**SD**		
Intercept | subject	76.6	8.76			0.88	0.94		
noise level [0 dB] | subject					1.29	1.14		
noise level [-6 dB] | subject					2.72	1.65		

noise level [Bab] | subject	4.32	2.08
	Number of obs.: 217, Subjects: 31	Number of obs.: 238, Subjects: 34

$SPIN\ score \sim 1 + noise\ level + (1|subject)$

$Intelligibility\ Rating \sim 1 + noise\ level + age + (1 + noise\ level|subject)$

Table S2. LMEM summary table for SPIN score vs Intelligibility score

	SPIN score			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***
(Intercept)	12.08	4.22	2.9	0.005
Intelligibility rating	8.2	0.58	14.2	**<0.001**
Random Effects		**Variance**		**SD**
Intercept | subject	71.98	8.49		
Number of obs.: 217, Subjects: 31				

$SPIN\ score \sim Intelligibility\ rating\ + (1|subject)$

Table S3. LMEM summary table for attended speech envelope reconstruction accuracies

	Reconstruction accuracy			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***
Intercept [Younger, Quiet]	0.30	0.01	47.9	**<0.001**
age [Older]	0.05	0.01	5.4	**<0.001**
noise level [0 dB]	-0.02	0.01	-3.2	**0.002**
noise level [-6 dB]	-0.01	0.01	-2.7	**0.007**
noise level [Bab]	-0.04	0.01	-6.9	**<0.001**
age [Older] × noise level [0 dB]	-0.02	0.01	-2.4	**0.018**
age [Older] × noise level [-6 dB]	-0.02	0.01	-2.1	**0.038**
age [Older] × noise level [Bab]	-0.02	0.01	-2.5	**0.014**
Random Effects	**Variance**		**SD**	
Intercept | subject	0.001		0.02	
Number of obs.: 782, Subjects: 34				

$Rec.\ accuracy \sim age \times noise\ level + (noise\ level|subject)$

Table S4. LMEM summary table for attended vs unattended speech envelope reconstruction accuracies

	Reconstruction accuracy			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***
Intercept [Attended, Younger]	0.29	0.01	49.9	**<0.001**
attention [Unattended]	-0.03	0.00	-6.9	**<0.001**
age [Older]	0.03	0.01	4.1	**<0.001**
Random Effects	**Variance**	**SD**		
Intercept | subject	0.0007	0.03		
attention | subject	0.0003	0.02		
	Number of obs.: 816, Subjects: 34			

$Rec. accuracy \sim age + attention + (1 + attention|subject)$

Table S5.LMEM summary table for attended TRF peak amplitudes

	$M50_{TRF}$ amplitude				$M100_{TRF}$ amplitude				$M200_{TRF}$ amplitude			
Fixed Effects	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***
Intercept [Younger, Quiet]	0.02	0.003	5.6	**<0.001**	0.02	0.004	5.2	**<0.001**	0.01	0.003	2.3	**0.021**
age [Older]	0.02	0.01	3.5	**<0.001**	0.01	0.01	0.9	0.369	0.03	0.004	6.8	**<0.001**
noise level [0 dB]	-0.01	0.003	-1.6	0.101	-0.00	0.003	-0.6	0.567	-0.00	0.003	-0.1	0.931
noise level [-6 dB]	-0.01	0.003	-2.1	**0.038**	-0.00	0.003	-0.3	0.745	-0.00	0.003	-0.5	0.612
noise level [Bab]	-0.02	0.003	-5.2	**<0.001**	-0.01	0.003	-2.6	**0.008**	-0.00	0.003	-1.1	0.289
age [Older] × noise level [0 dB]	-0.01	0.005	-2.1	**0.034**	0.01	0.005	2.9	**0.004**	-0.02	0.004	-4.8	**<0.001**
age [Older] × noise level [-6 dB]	-0.01	0.005	-2.9	**0.004**	0.01	0.005	2.6	**0.009**	-0.02	0.004	-4.9	**<0.001**
age [Older] × noise level [Bab]	-0.01	0.005	-2.2	**0.029**	0.01	0.005	2.1	**0.036**	-0.03	0.004	-5.8	**<0.001**
Random Effects	**Variance**	**SD**			**Variance**	**SD**			**Variance**	**SD**		
Intercept | subject	0.0001	0.01			0.0002	0.01			0.0001	0.01		
	Number of obs.: 136, Subjects: 34				Number of obs.: 136, Subjects: 34				Number of obs.: 136, Subjects: 34			

$M50_{TRF} \sim age \times noise\ level + (1|subject)$

$M100_{TRF} \sim age \times noise\ level + (1|subject)$

$M200_{TRF} \sim age \times noise\ level + (1|subject)$

Table S6. LMEM summary table for attended vs unattended TRF peak amplitudes

	$M50_{TRF}$ amplitude				**$M100_{TRF}$ amplitude**				**$M200_{TRF}$ amplitude**			
Fixed Effects	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***
Intercept [Attended, Younger, 0 dB]	0.02	0.000	7.6	**<0.001**	0.02	0.000	6.1	**<0.001**	0.01	0.000	4.8	**<0.001**
attention [Unattended]	0.01	0.00	5.8	**<0.001**	-0.01	0.00	-5.9	**<0.001**	-0.01	0.00	-4.7	**<0.001**
age [Older]					0.02	0.00	3.8	**<0.001**	0.01	0.00	3.7	**<0.001**
attention [Unattended] × age [Older]					-0.01	0.00	-4.1	**<0.001**				
noise level [-6 dB]									-0.00	0.00	-1.0	0.310
attention [Unattended] × noise level [-6 dB]									0.00	0.00	1.2	0.237
Random Effects	**Variance**	**SD**			**Variance**	**SD**			**Variance**	**SD**		
Intercept | subject	0.0001	0.01			0.0001	0.01			0.0001	0.00		
	Number of obs.: 136, Subjects: 34				Number of obs.: 136, Subjects: 34				Number of obs.: 136, Subjects: 34			

$M50_{TRF} \sim attention + (1|subject)$

$M100_{TRF} \sim age \times attention + (1|subject)$

$M200_{TRF} \sim age \times noise\ level + (1|subject)$

Table S7. LMEM summary table for attended TRF peak latencies

	$\mathbf{M50_{TRF}}$ **Latency**				$\mathbf{M100_{TRF}}$ **Latency**				$\mathbf{M200_{TRF}}$ **Latency**			
Fixed Effects	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***	***Est.***	***SE***	***t value***	***p value***
Intercept [Younger, Quiet]	42.44	2.85	14.9	**<0.001**	120.57	3.32	36.3	**<0.001**	178.70	5.20	34.4	**<0.001**
age [Older]	-2.27	4.11	-0.6	0.581	-9.44	4.84	-1.9	0.051	11.05	7.22	1.5	0.126
noise level [0]	18.48	3.02	6.1	**<0.001**	14.59	2.75	5.3	**<0.001**	21.80	4.64	4.7	**<0.001**
noise level [-6]	13.07	2.99	4.4	**<0.001**	17.65	2.75	6.4	**<0.001**	29.97	4.91	6.1	**<0.001**
noise level [Bab]	25.63	3.12	8.2	**<0.001**	16.99	2.73	6.2	**<0.001**	30.82	5.12	6.0	**<0.001**
noise level [0] × age [Older]	-7.66	4.32	-1.8	0.076	8.05	3.98	2.0	**0.043**	17.08	6.68	2.6	**0.011**
noise level [-6] × age [Older]	-0.49	4.30	-0.1	0.909	11.23	3.97	2.8	**0.004**	11.16	6.97	1.6	0.109
noise level [Bab] × age [Older]	-16.39	4.54	-3.6	**<0.001**	8.64	3.95	2.2	**0.023**	-6.92	7.11	-0.9	0.330
Random Effects	**Variance**	**SD**			**Variance**	**SD**			**Variance**	**SD**		
Intercept | subject	60.9	7.8			128.2	11.32			270.8.	16.46		
	Number of obs.: 125, Subjects: 34				Number of obs.: 131, Subjects: 34				Number of obs.: 109, Subjects: 32			

$M50_{TRF} \sim age \times noise\ level + (1|subject)$

$M100_{TRF} \sim age \times noise\ level + (1|subject)$

$M200_{TRF} \sim age \times noise\ level + (1|subject)$

Table S8. LMEM summary table for attended vs unattended TRF peak latencies

	$\mathbf{M50_{TRF}}$ **Latency**				$\mathbf{M100_{TRF}}$ **Latency**				$\mathbf{M200_{TRF}}$ **Latency**			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***	***Estimates***	***SE***	***t value***	***p value***	***Estimates***	***SE***	***t value***	***p value***
Intercept [Younger, Attended]	60.76	2.79	21.8	**<0.001**	136.62	2.73	50.1	**<0.001**	202.15	4.90	41.3	**<0.001**
age [Older]	-9.76	4.03	-2.4	**0.015**	-1.97	3.93	-0.5	0.616	27.84	7.09	3.9	**<0.001**
attention [Unattended]	0.50	3.11	0.2	0.873	8.03	3.21	2.5	**0.012**	-12.02	6.28	-1.9	0.056
noise level [-6 dB]	-5.40	2.17	-2.5	**0.013**	2.95	1.46	2.0	**0.043**	7.15	2.38	3.0	**0.003**
age [Older] × attention [Unattended]	-2.25	4.48	-0.5	0.616	-12.01	4.56	-2.6	**0.008**	-28.54	8.26	-3.5	**0.001**
attention [Unattended] × noise level [-6 dB]	6.80	3.03	2.3	**0.025**					-12.04	3.96	-3.0	**0.002**
age [Older] × noise level [-6 dB]	7.15	3.08	2.3	**0.020**								
age [Older] × attention [Unattended] × noise level [-6 dB]	-9.05	4.33	-2.1	**0.037**								
Random Effects	**Variance**	**SD**			**Variance**	**SD**			**Variance**	**SD**		
Intercept | subject	97.58	9.88			101.24	10.06			330.42	18.18		
attention [Unattended] | subject	89.56	9.43			117.18	10.82			313.69	17.71		
noise level [-6 dB] |					31.28	5.59						

subject

	Number of obs.: 133, Subjects: 34	Number of obs.: 116, Subjects: 34	Number of obs.: 94, Subjects: 32

$M50_{TRF} \sim age \times attention \times noise\ level + (1 + attention|Subject)$

$M100_{TRF} \sim age \times attention + noise\ level + (1 + attention + noise\ level|subject)$

$M200_{TRF} \sim age \times attention + attention \times noise\ level + (1 + attention|subject)$

Table S9. LMEM summary table for $M200_{TRF}$ amplitudes vs latency

	$M200_{TRF}$ amplitude			
Fixed Effects	***Estimates***	***SE***	***t value***	***p value***
Intercept [Younger]	-0.01	0.02	-0.6	0.585
age [Older]	0.08	0.02	3.5	**<0.001**
latency	0.00008	0.00009	0.9	0.355
age [Older] × latency	-0.0003	0.0001	-2.9	**0.004**
Random Effects	**Variance**	**SD**		
Intercept | subject	0.0004	0.006		
	Number of obs.: 55, Subjects: 29			

$$M200_{TRF}amplitude \sim age \times M200_{TRF}latency + (1|subject)$$

Table S10. GAMM summary table for Integration Window Analysis

	Reconstruction Accuracy			
Parametric Coefficients	***Estimates***	***SE***	***t value***	***p value***
Intercept [Younger, Quiet]	0.25	0.00	55.42	**<0.001**
Older × Quiet	0.04	0.01	6.43	**<0.001**
Younger × 0 dB × Attended	-0.02	0.001	-4.56	**<0.001**
Older × 0 dB × Attended	0.01	0.01	1.45	0.146
Younger × 0 dB × Unattended	-0.03	0.001	-8.21	**<0.001**
Older × 0 dB × Unattended	-0.01	0.01	-1.78	0.075
Younger × -6 dB × Attended	-0.01	0.001	-3.50	**<0.001**
Older × -6 dB × Attended	0.02	0.01	2.29	**0.022**
Younger × -6 dB × Unattended	-0.03	0.001	-8.19	**<0.001**
Older × -6 dB × Unattended	-0.00	0.01	-0.11	0.914
Smooth Terms	**edf**		**Ref.df**	***p***
s(Wind) :Younger × Quiet	5.93		370.86	**<0.001**
s(Wind) :Older × Quiet	5.94		520.27	**<0.001**
s(Wind) :Younger × 0 dB × Attended	5.80		243.32	**<0.001**
s(Wind) :Older × 0 dB × Attended	5.90		336.61	**<0.001**
s(Wind) :Younger × 0 dB × Unattended	5.89		231.45	**<0.001**
s(Wind) :Older × 0 dB × Unattended	5.94		298.78	**<0.001**
s(Wind) :Younger × -6 dB × Attended	5.82		249.48	**<0.001**

s(Wind) :Older × -6 dB × Attended	5.86	330.57	**<0.001**
s(Wind) :Younger × -6 dB × Unattended	5.91	227.02	**<0.001**
s(Wind) :Older × -6 dB × Unattended	5.94	302.23	**<0.001**
s (Wind,Subject)	107.49	5.68	**<0.001**

Number of obs.: 14144, Subjects: 34