Linguistic speech processing decreases with advancing age

Marlies Gillis[1A] and Jill Kries[1A], Maaike Vandermosten [1B] and Tom Francart[1B]

[1]: Experimental Oto-Rhino-Laryngology, Department of Neurosciences, Leuven Brain Institute, KU Leuven, Belgium

[A]: Shared first authorship

[B]: Shared last authorship

Corresponding authors:

Marlies Gillis (marlies.gillis@kuleuven.be),

Jill Kries (jill.kries@kuleuven.be)

Abstract

Background: Older adults process speech differently, but it is not yet clear how aging affects different levels of processing natural, continuous speech, both in terms of bottom-up acoustic analysis and top-down generation of linguistic-based predictions. We studied natural speech processing across the adult lifespan via electroencephalography (EEG) measurements of neural tracking.

Goals: Our goals are to analyze the unique contribution of linguistic speech processing across the adult lifespan using natural speech, while controlling for the influence of acoustic processing. In particular, we focus on changes in spatial and temporal activation patterns in response to natural speech across the lifespan.

Methods: 52 normal-hearing adults between 17 and 82 years of age listened to a naturally spoken story while the EEG signal was recorded. We investigated the effect of age on acoustic and linguistic processing of speech. Because age correlated with hearing capacity and measures of cognition, we investigated whether the observed age effect is mediated by these factors. Furthermore, we investigated whether there is an effect of age on hemisphere lateralization and on spatiotemporal patterns of the neural responses. **Results:** Our EEG results showed that linguistic speech processing declines with advancing age. Moreover, as age increased, the neural response latency to certain aspects of linguistic speech processing increased. Also acoustic neural tracking (NT) decreased with increasing age but in contrast to linguistic processing, older subjects showed shorter latencies for early acoustic responses to speech. No evidence was found for hemispheric lateralization in neither younger nor older adults during linguistic speech processing. Most of the observed aging effects on acoustic and linguistic processing were not explained by age-related decline in hearing capacity or cognition. However, our results suggest that the effect of increasing neural response latency with age for word-level linguistic processing is likely more due to an age-related decline in cognition than a robust effect of age.

Conclusion: Spatial and temporal characteristics of the neural responses to continuous speech change across the adult lifespan for both acoustic and linguistic speech processing. These are most likely explained by compensation mechanisms that occur during healthy aging. Traces of this compensation mechanism are reported as modifications in the neural response to acoustic and linguistic speech characteristics in terms of spatial and temporal differences, suggesting structural and functional changes. Additionally, different frontal patterns were observed for linguistic processing, suggesting that older adults recruit additional frontal brain regions to compensate for the age-related decline.

Keywords: neural tracking, speech processing, lifespan, aging, linguistic processing, speech understanding

Highlights:

- With increasing age, linguistic neural tracking of speech decreases.
- With increasing age, the processing speed of linguistic aspects of speech slows down.
- The processing speed for word-level linguistics is likely affected by cognition.
- Older adults rely on compensation mechanisms during speech processing.
- The compensation mechanism affects spatial and temporal aspects of neural responses.

1 Introduction

Aging goes hand in hand with physical and behavioral change. The brain is not spared and undergoes micro- and macrobiological alterations in structure and function. Processing speech and deducing meaning from it is a complex task that the human brain manages in a continuous manner at a fast pace. At an older age, speech comprehension seems to be largely preserved (Burke and Shafto, 2008; Shafto and Tyler, 2014), at least in optimal conditions, yet using EEG, researchers have found less efficient speech processing in older adults (Federmeier et al., 2003) and evidence for differential recruitment of neural resources (Brodbeck et al., 2018b). So far, it has not been investigated whether the observed age effects on speech processing are driven by differences in lower-level acoustic and/or higher-level linguistic speech processing. In this study, we explore the effects of age on neural resource allocation for natural, continuous speech processing using EEG-based NT. NT is the phenomenon whereby the brain time-locks to specific aspects of the incoming speech (Brodbeck and Simon, 2020, for a review,). This method allows to explore aging effects separately for bottom-up and top-down aspects of speech processing.

1.1 Age-related changes in acoustic speech processing

The acoustic analysis in the primary auditory cortex consists of analyzing fundamental frequency, formant transitions and amplitude modulations in the speech envelope. By using EEG or magnetoencephalography (MEG), the neural response to auditory stimuli can be examined by averaging the signal over a large number of repetitive stimuli (i.e., event-related potentials (ERP)). Cortical auditory evoked potentials (CAEP), such as the P1-N1-P2 complex, reflect sound detection and encoding in the auditory cortex (Martin et al., 2008; Harris, 2020). In clinically normal-hearing older adults, CAEP peak amplitudes and latencies (N1 and P2) are increased compared to younger adults (McCullagh and Shinn, 2013; Tremblay et al., 2002, 2003). This has been related to decreased behavioral results in older adults during a phoneme discrimination task (Tremblay et al., 2002, 2003). Evidence also suggests that older adults may use different neural resources than younger adults for the acoustic analysis. More specifically, younger adults typically display a right preference of the primary auditory cortex for processing slow modulation frequencies at 4 Hz (i.e., a modulation frequency that coincides with the rate of syllables in speech), yet older adults show a greater reliance on the left hemisphere (Farahani et al., 2020; Goossens et al., 2016). These findings are based on artificially created experimental stimuli (e.g., repetition of isolated linguistic units) and therefore not an ideal approximation of real-life communication. NT of connected speech presents a novel way to analyze neural responses in more natural listening conditions. When using the envelope as speech representation, increased neural tracking has been found in older adults compared to younger adults in response to a narrative (Presacco et al., 2016; Brodbeck et al., 2018b; Decruy et al., 2019). The envelope consists of the slow amplitude modulations of speech. Envelope tracking mainly

characterizes lower-level acoustic processing, but it has also been shown to be modulated by speech intelligibility and higher-order speech-specific processing (Peelle et al., 2013; Vanthornhout et al., 2018; Prinsloo and Lalor, 2020). As of yet, it is not clear whether the observed age differences in envelope tracking are driven by processing of acoustics or of higher-order aspects. Brodbeck et al. (2018b) found that enhanced envelope tracking in older adults arises early on in areas inferior and lateral to the left primary auditory cortex. However, it is not yet clear how this result can be interpreted in terms of underlying functions. Put differently, which aspects of speech do older adults process differently than younger adults, acoustic and/or linguistic aspects?

1.2 Age-related changes in linguistic speech processing

To assign meaning to incoming speech, neural signals related to bottom-up acoustic cues (e.g., phoneme and word onsets) interact in a parallel fashion with top-down processes (e.g., generating predictions about linguistic units) (Federmeier, 2007; Kuperberg and Jaeger, 2016; Brodbeck et al., 2021; Gwilliams and Davis, 2022; Hamilton et al., 2021). ERP research has found the N400 peak to be a neural correlate of semantic information, specifically its activation and its integration into the sentence context (Kutas and Federmeier, 2011; Nieuwland et al., 2020). This negative N400 response usually appears between 200 and 600 ms and is largest over centro-parietal electrodes (Kutas and Federmeier, 2011). In older adults, studies have consistently reported decreased N400 amplitudes and temporally delayed responses compared to younger adults (Kutas and Iragui, 1998; Woodward et al., 1993; Federmeier and Kutas, 2005; Tiedt et al., 2020; Gunter et al., 1992; Federmeier et al., 2002). This effect has been ascribed to a declining efficiency in older adults to generate predictions based on sentence-level context information (i.e. if a word is semantically fitting the previous words in the sentence) (Tiedt et al., 2020). However, this finding is difficult to unify with behavioral results, which generally report preserved or increased semantic and world knowledge in the elderly (Shafto and Tyler, 2014; Spreng and Turner, 2019) as well as a greater sensitivity to linguistic context (Tiedt et al., 2020). A possible explanation for these rather conflicting results may be that elderly make use of different strategies and neural structures to behaviorally maintain a viable level (Tiedt et al., 2020; Reuter-Lorenz and Cappell, 2008). Indeed, evidence suggests that older adults present a more bilaterally organized activation pattern for processing linguistic aspects of speech, whereas younger adults rely more on the left hemisphere (Diaz et al., 2016; Wlotko et al., 2010; Wierenga et al., 2008; Grossman et al., 2002; Wingfield and Grossman, 2006). This finding is yet to be tested via EEG-based NT.

Although EEG research using naturally spoken stories has traditionally focused on bottom-up processing aspects of speech, recent NT studies demonstrate its potential to gain insights into higher-order language processing by analyzing the neural response related to the amount of linguistic information conveyed by a phoneme or word (Brodbeck et al., 2018a; Alday, 2019; Brennan and Hale, 2019; Weissbart et al., 2019;

Broderick et al., 2021; Mesik et al., 2021; Shain et al., 2020; Gillis et al., 2021b). The neural response to the surprisal of a word or phoneme represents the mismatch between top-down predicted and bottom-up perceived linguistic units. Hence, the more unexpected or surprising the linguistic unit is, the larger the prediction error is (Weissbart et al., 2019; Friston, 2010). The relationship between such a surprisal model and the EEG signal can be studied via NT. This way, Broderick et al. (2021) found longer latencies of the neural response to word surprisal in older adults. However, the study by Mesik et al. (2021) did not confirm these results of Broderick et al. (2021). In both studies, the negative peak appeared during the same time window as the N400 ERP and displayed a similar centro-parietal topography. From functional magneto-resonance imaging (fMRI) research measuring the hemodynamic activity in response to a story, the word surprisal peak seems to be specific to the language network (Shain et al., 2020). These EEG studies investigating linguistic speech processing in older adults did not control the influence of acoustic processing in their model and therefore, the results may be biased by age-related effects in bottom-up acoustic processing (Gillis et al., 2021b).

1.3 Current study

In this study, we investigated age-related spatiotemporal neural changes in acoustic and linguistic aspects of speech processing respectively. Furthermore, we analyzed for the first time the unique contribution of linguistic speech processing in elderly by controlling for acoustic and lexical segmentation representations (Gillis et al., 2021b). To do so, we investigated EEG responses to natural, connected speech in neurologically healthy, normal-hearing adults ranging from 17 to 82 years of age. We investigated neural tracking of speech using a forward modeling approach. This method describes how and how well the brain responds to speech representations in a time-locked fashion, respectively by the temporal response function (TRF) ("how") and prediction accuracy ("how well"). The prediction accuracy is a measure of neural tracking (NT), i.e., the better the brain tracks the incoming speech, the higher the prediction accuracy will be. In order to investigate the unique contribution of linguistic speech processing, we analysed the increase in prediction accuracy when linguistic representations are added to a model consisting of acoustic and lexical segmentation representations. This increase in prediction accuracy, which can be attributed to the linguistic representation, is further denoted as linguistic tracking, i.e., the tracking of linguistic aspects of the incoming speech when controlling for acoustics and lexical segmentation. We here solely focused on linguistic aspects related to generating predictions about upcoming speech.

For acoustic and linguistic speech processing separately, we analyzed the effect of age on the magnitude of NT. Given that age, hearing and cognition are interrelated, we performed additional analyses to examine whether any observed age effects on the acoustic and linguistic aspects of speech are mediated by decline in hearing or cognitive performance respectively (1a). Spatial pattern differences across age were investigated using NT outcomes in predefined regions of interest and in the hemispheres (1b). Temporal

pattern changes across age were analyzed drawing on TRF peak latencies. Again, we investigated whether the effect of age on the latency was mediated by hearing capacity or cognition (2a). Additionally, the topographies of the TRF peaks were investigated (2b).

For acoustic speech processing, we hypothesized to find increased NT with increasing age (Presacco et al., 2016; Brodbeck et al., 2018b; Decruy et al., 2019). This effect was expected to be strongest in the auditory cortex, situated in temporal regions of the brain (Cardin, 2016). Moreover, based on previous findings, we hypothesized that there would be stronger left-hemispheric lateralization in older than in younger adults for the acoustic model (Brodbeck et al., 2018b). Regarding temporal patterns, we hypothesized longer latencies with advancing age, given that most studies investigating CAEP reported longer latencies in older adults. Additionally, we hypothesized that this age effect is driven by hearing capacity: with increasing age, hearing capacity decreases, which might affect neural speech processing. Previous literature showed that adults with hearing loss showed higher acoustic NT and later neural response latencies compared to normal hearing adults (Gillis et al., 2021a), which supports our hypothesis.

For linguistic speech processing, we hypothesized a decrease in linguistic NT with advancing age. Although Broderick et al. (2021) and Mesik et al. (2021) reported diverging results comparing linguistic NT between younger and older adults, ERP studies consistently suggested that older adults show smaller and later N400 responses. Although the link between NT and ERP amplitudes is not well established, we assume that language processing changes across the lifespan, which would affect both the NT as well as the spatial and temporal characteristics of the neural responses. Based on ERP (Kutas and Federmeier, 2011) and NT (Broderick et al., 2021; Mesik et al., 2021) literature, we expected the age-effect to be strongest over the centro-parietal region. Concerning hemispheric lateralization, we hypothesized more symmetrical NT in older than in younger adults for the linguistic speech processing (Diaz et al., 2016; Wlotko et al., 2010). Regarding temporal dynamics, the TRF latencies of linguistic speech representations were expected to be longer in older adults, based on the results from Broderick et al. (2021), who found increased latencies in older adults in response to word surprisal. Moreover, we hypothesized that the effect of age is mediated by cognitive factors: with increasing age, the brain undergoes structural changes which goes hand in hand with a decline in cognition. Structural changes in the brain imply changes in the spatial and temporal dynamics of the neural responses. Therefore, we hypothesized that cognition would impact linguistic NT and the neural response latencies to linguistic representations.

2 Material and Methods

We investigated acoustic and linguistic speech processing across the adult life span. For each type of speech processing, we investigated two outcome variables: *the prediction accuracy*, also denoted as NT, to investigate age-trends and lateralization across age, and the *TRF pattern* to study how the spatial and temporal characteristics of the neural responses to the speech representations change with age. We used a subset of the data collected in the study by Decruy et al. (2019). In comparison to Decruy et al. (2019), we investigated NT of a more extensive set of speech representations covering acoustic and linguistic aspects of the speech, rather than just the speech envelope investigated in Decruy et al. (2019).

2.1 Participants

We analyzed the data of 52 normal-hearing adults. Age varied between 17 and 82 years old. All participants had a normal hearing, confirmed with pure tone audiometry (no hearing threshold exceeded 30 dB HL at all octave frequencies from 125 to 4000 Hz; defined identical to Decruy et al. (2019)). The average of these measured hearing thresholds is denoted as pure tone average (PTA) (octave frequencies from 125 to 4000 Hz; mean$\pm$std = 12.1$\pm$5.08; the distribution of PTA across age is visualized in fig. 2). Same as in Decruy et al. (2019), we excluded persons with a low cognitive score, i.e., Montreal Cognitive Assessment (Nasreddine, 2004) $<$ 26/30, and persons with a learning disability or medical history of head concussions. The data of two participants of the original dataset, collected by Decruy et al. (2019), were discarded: one participant was discarded because the stimulus was presented to the left instead of the right ear; another participant was missing an EEG electrode. The experiments were carried out following a protocol accepted by the KU Leuven Ethical Committee. All participants signed an informed consent form. For more details regarding the cognitive screenings, we refer to Decruy et al. (2019).

2.2 Cognitive testing

Decruy et al. (2019) conducted two cognitive tests. Cognitive performance was assessed using the Stroop color-word test (SCWT) (Hammes, 1978), measuring inhibition, and the reading span test (RST), measuring working memory (van den Noort et al., 2008; Vercammen et al., 2017).

During the RST, the participants were instructed to read sets of 20 sentences aloud. After each set, the participants recalled as many as possible final words of sentences of the set. After three sets of sentences, the working capacity is represented by the amount of correctly recalled final words of 60 sentences as possible (resulting in a score of 60). A higher RST outcome represents better working memory performance.

For the SCWT, participants were shown three cards containing respectively color names, colored rectangles and color names printed in incongruent colors. For each card, the task instruction was to name or

read the color shown on the card as fast as possible. For the card containing the color names printed in incongruent colors, the task consisted of naming the color of the word and inhibiting to read the color name. The SCWT score is calculated as the difference in response time (in milliseconds) between the third and the second testing condition. Therefore, a higher SCWT score represents a higher difference in response time and thus a decreased inhibition performance.

For a more extensive explanation of how the cognitive tests were carried out, we refer to Decruy et al. (2019).

2.3 EEG testing

During the EEG experiment, the participants listened to a story presented in silence. This story lasted 12 minutes and was narrated by a male speaker.

The stimuli were presented to the right ear through at an intensity of 55 dB SPL (A-weighted). During the stimulus presentation, EEG data was collected using a BioSemi Active Two system (Amsterdam, Netherlands) with 64 EEG electrodes. For more details on the experimental protocol, we refer to Decruy et al. (2019).

2.4 EEG processing

Firstly, to decrease the processing time, the EEG recording with a sampling frequency of 8192 Hz was downsampled to 256 Hz. Subsequently, we filtered the EEG using a multi-channel Wiener filter (Somers et al., 2018) to remove artifacts due to eye blinks and referenced the EEG to the common-average. We filtered the data between 0.5 and 25 Hz using a Chebyshev filter (Type II with an attenuation of 80 dB at 10% outside the passband). Lastly, an additional downsampling step was done to reduce the sampling rate to 128 Hz.

2.5 Predictors

The speech representations used in this study are based on the findings of Gillis et al. (2021b). We divided the speech representations into two categories (1) acoustic representations, and (2) linguistic speech representations. The acoustic representations resembled the continuous acoustic energy of the presented speech material, while the linguistic speech representations denote to what extent a phoneme or word can be predicted based on the context.

The acoustic speech representations were calculated from the speech stimulus. First, to account for the filtering of the ER-3A insert earphones, the speech stimulus was low-pass filtered below 4000 Hz (zero-phase low-pass FIR filter with a hamming window of 159 samples). The spectrogram representation was calculated using the Gammatone Filterbank Toolkit 1.0 (Heeris, 2014) with center frequencies between 70 and 4000 Hz with 256 filter electrodes and an integration window of 0.01 second. This toolkit was used

to calculate a spectrogram representation based on a series of gammatone filters inspired by the human auditory system's structure (Slaney, 1998). The resulting 256 filter outputs were averaged into eight frequency bands (each containing 32 gammatone filter outputs). Additionally, each frequency band was downsampled to the same sampling frequency as the processed EEG, namely 128 Hz. Subsequently, the acoustic onsets representation was calculated as the half-wave rectification of the spectrogram's derivative.

Linguistic speech representations were obtained at the phoneme and word level. ***Phoneme surprisal*** and ***cohort entropy*** were used as two linguistic speech representations at the phoneme level (Brodbeck et al., 2018a). Both representations were derived from the activated cohort of words, given a certain sound (e.g., when hearing the sound *sp*, the activated cohort of words include words like *speed*, *spam*, *special*, etc.). Phoneme surprisal represents the surprisal of a phoneme. It was calculated as the negative logarithm of the phoneme probability in the activated cohort. In contrast, cohort entropy reflects the degree of competition between the words congruent with the acoustic input at a given point during a word. It was calculated as the Shannon entropy of the words in the active cohort. A higher value indicated that more phonemes have a high probability of occurring as a subsequent phoneme. Gillis et al. (2021b) showed that the EEG response to these representations is characterized by a topography with central negativity around 250 ms. Concerning linguistic representations at the word level, ***Word surprisal*** was calculated, which reflects the surprisal of a word given the four previous words. It was calculated as the negative logarithm of the conditional probability of the word, given the four preceding words. The response to word surprisal is characterized by the topography with central negativity around 400 ms (Weissbart et al., 2019; Gillis et al., 2021b). This response resembles the N400-effect in ERP responses.

To control for the early acoustic response, we included speech representations for phoneme and word onsets, i.e., lexical segmentation. The phoneme and word onset timings were determined by forced alignment of the identified phonemes using the speech alignment component of the reading tutor (Duchateau et al., 2009). The phoneme and word onset representations describe the onsets with a hot-one encoding: 1 at the time of onset while all other values are 0. As the response to phoneme and word onsets is not purely acoustic, we did not include these representations when evaluating acoustic NT. When investigating linguistic tracking of speech, controlling for acoustic and lexical segmentation representations is necessary (Gillis et al., 2021b). If this step is not considered, a linguistic representations might be significantly tracked due to correlations with acoustic or lexical segmentation representations.

2.6 Neural tracking

The forward modelling approach provides insight into the brain's response to speech in two ways: a measure of NT, the prediction accuracy, and a temporal response function (TRF), which characterizes the brain's temporal response to the speech representations. The TRF models how the brain responds

to these speech representations. Subsequently, the TRF is used to predict the EEG by convolving it with the speech representations. Correlating the predicted responses with the measured EEG results in a prediction accuracy for each EEG electrode. These prediction accuracies are a measure of NT: the higher the prediction accuracy, the more the brain tracks the speech representations.

We used the Eelbrain toolbox to estimate the TRF and to calculate the prediction accuracies. We estimated the TRF using boosting (David et al., 2007), in a cross-validation approach with ten partitions and an integration window between 0 and 600 ms (using selective stopping based on the ℓ_2-norm and Hamming window of 50 ms as basis function). Additionally, we calculated the prediction accuracy in three different testing windows that each characterize a different type of response:

- 30 to 160 ms: the early acoustic response like N1 in the spectrogram
- 160 to 300 ms: a later acoustic response like the P2 in the spectrogram and phoneme recognition which occurs around 250 ms as observed in the TRF to phoneme surprisal and cohort entropy
- 300 to 600 ms: linguistic processes like the N400 response to word surprisal

These testing windows were determined based on the average TRF across all participants. Outside the testing windows, we set the TRF to zero. This adapted TRF was convolved with the speech representations to obtain the prediction accuracy based on the lags specified in the testing window. We opted to set the TRF, trained on the integration window from 0 to 600 ms, partially to zero over retraining the TRF in different integration windows. This was done to prevent edge effects, i.e., high values at the beginning and end of the window, which occur due to autocorrelations in the speech.

We classified NT into two different types: (1) acoustic NT and (2) linguistic NT. Acoustic NT denoted the prediction accuracies of a model using acoustic speech representations of the speech (i.e., spectrogram and acoustic onsets). Linguistic NT was calculated as the prediction accuracy increase when linguistic speech representations are included in the model on top of acoustic speech representations and lexical segmentation representations as phoneme and word onsets. In more detail, we subtracted prediction accuracies from the model that included acoustic and lexical segmentation representations from those obtained with the model that included acoustic, lexical segmentation, and linguistic speech representations. The difference between these models thus constitutes the added value of linguistic representations over and beyond acoustic and lexical segmentation properties. Therefore, if the prediction accuracy is significantly higher than zero, the linguistic speech representations explain unique information over and beyond the response to the acoustic signal or the onset of a word or phoneme. If acoustic and lexical segmentation representation are not taken into account, spurious tracking might be observed leading to biased results.

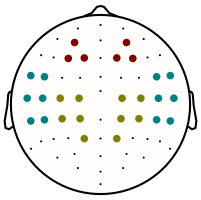

Figure 1: Different electrodes selections: frontal (red), centro-parietal (green) and temporal (blue).

2.7 Regions of interest

We investigated the effect of age on acoustic NT and linguistic NT averaged across all EEG electrodes. However, this approach can average out sublte and local effects. Therefore, we also investigated acoustic NT and linguistic NT in three a priori-chosen electrode selections, i.e., a frontal electrode selection, a centro-parietal electrode selection, and a temporal electrode selection (visualized in fig. 1).

2.8 Determination of lateralization

To determine whether there was an age-dependent hemisphere effect, we averaged the prediction accuracy across respectively right or left-side electrodes, excluding central electrodes. This analysis was performed across all EEG electrodes and for the individual ROIs.

2.9 Determination of peak latency

To determine the peak latency, we defined a time window for each peak of interest, based on the average TRF across all participants, wherein we investigated the amplitude and latency. The resulting time windows are summarized in table 1. We broadened these time windows on each end with 30 ms to generalize these regions across all participants because the average TRF is biased towards 50-year-olds as age was not equally distributed across all participants (see age-distribution in fig. 2). Within each time window, we determined the root mean square (RMS) of the TRF across electrodes for each time lag. The RMS is higher when many electrodes have high activity, indicating the possibility of a peak. In each local maximum of the RMS, we correlated the peak topography with a template peak topography, i.e., a peak's topography determined across all participants and averaged across the whole time window (table 1). We selected the latency of the local maximum where the peak topography had the highest correlation with the template peak topography. Additionally, we set a criterion that more than 30 electrodes must have the same sign as the template peak topography. If not, we discarded the peak latency and topography from the analysis. The number of peaks found per speech representation and time window is shown in table 2. Code for this method to determine peak latencies is available online (https://github.com/MarliesG/TemplateMatching).

Table 1: Selected time windows for each speech representation

Speech representations	**category**	**time windows**
spectrogram	acoustic	30 - 125 ms 150 - 250 ms
acoustic onsets	acoustic	30 - 70 ms 100 - 150 ms
phoneme surprisal	linguistic	80 - 150 ms 250 - 300 ms
cohort entropy	linguistic	100 - 175 ms 250 - 350 ms
word surprisal	linguistic	125 - 225 ms 350 - 450 ms

Table 2: Percentage of peaks found per speech representation and time window

Speech representation	**age-label**	**early window**	**late window**
spectrogram	younger	93%	100%
	middle	95%	100%
	older	100%	100%
acoustic onsets	younger	100%	100%
	middle	100%	86%
	older	93%	100%
phoneme surprisal	younger	93%	100%
	middle	86%	95%
	older	100%	93%
cohort entropy	younger	87%	100%
	middle	86%	100%
	older	100%	87%
word surprisal	younger	87%	100%
	middle	100%	100%
	older	100%	87%

In total 52 participants: 15 younger adults, 22 middle aged adults, 15 older adults

2.10 Statistics

The statistical analyses were carried out with Rstudio (RStudio Team, 2020), which uses R software (R Core Team, 2020). All linear models to determine age effects on the prediction accuracy, averaged across all electrodes, or peak latency were created using the lme R package (Bates et al., 2015). The model characteristics are reported with the adjusted R^2_{adj}-value, F-value, degrees of freedom and p-value. The estimates of the considered effect are reported with estimates, represented by β, standard errors, t-values and corresponding p-values.

Similarly as the linear models, the linear mixed effect models were created using the lme R package (Bates et al., 2015). These models were used to investigate the effect of age on the prediction accuracy in specific ROI and/or hemipheric lateralization. To investigate the effect of age on the prediction accuracy in the different ROI, linear mixed effects model were created with ROI as a 3-level factor variable (fixed effect) and participant as random effect (*prediction accuracy∼age+ROI+age*ROI+(1|participant)*). To investigate the effect of age on lateralization, we determined whether the effect depended on the considered hemi-

sphere by using hemisphere as a 2-level factor, age as fixed effect, and participant as a random effect (*prediction accuracy∼age+hemisphere+age*hemisphere+(1|participant)*). In case the hemisphere and age interaction effect was significant across all electrodes, the same analysis was repeated separately for the 3 region of interest (ROI)s (*prediction accuracy in ROI∼age+hemisphere+age*hemisphere+(1|participant)*). false discovery rate (FDR) correction was applied to account for multiple comparisons (Benjamini and Hochberg, 1995).

These linear mixed effect models were fitted using using the restricted maximum likelihood method. Post-hoc pair-wise comparisons were performed using the R package emmeans (Lenth et al., 2022), the Tukey contrasts were applied and corrected for multiple comparisons applying the FDR method (Benjamini and Hochberg, 1995). The model characteristics are reported with the adjusted $R^2 conditional$-value. Similar to the R^2_{adj} for linear models, $R^2 conditional$ represents the estimator of the explained variance by the linear mixed effect model (Barton, 2022). The estimates of the considered effect are reported with estimates, represented by β, standard errors, degrees of freedom, t-values and corresponding p-values.

To examine whether the direct effect of age persists after considering the correlated behavioral measures, i.e., PTA, SCWT, and RST, we applied structural equation modelling (SEM). We applied SEM using the Lavaan R package (Rosseel, 2012). This package allows estimating the indirect effect of a certain predictor through other predictors. To assess the goodness of fit, we used the criteria according to Hu and Bentler (1999): Comparative Fit Index (CFI) > 0.95, Tucker-Lewis Index (TLI) > 0.95, root mean squared error of approximation (RMSEA) < 0.06. For this analysis, all variables were scaled by subtracting the mean and dividing by their standard deviations. The outcomes of SEM are reported as estimates of the effect, represented by β, the standard error (SE), z-value, and p-value.

To investigate the effect of age on the topographies and TRF patterns, we created two groups consisting of the 15 oldest and 15 youngest participants. Although we selected an arbitrary number of younger and older participants, this division into two groups allows us to compare how the TRF patterns or NT differed between younger and older participants. For investigating differences in topographies, we used two different methods. Firstly, an independent cluster-based permutation test proposed by Maris and Oostenveld (2007), using the Eelbrain implementation (Brodbeck, 2020). This method allows pinpointing the electrodes which drive the significant difference between both groups. However, a disadvantage of using such a test is a preset threshold for defining a cluster. Depending on the value of this threshold, the test is more specific to either weaker and widespread effects or strong and localized effects. To overcome this, we applied the method proposed by McCarthy and Wood (1985). This method compares normalized topographies, therefore discarding amplitudes effects. More specifically, the method is based on an ANOVA test for an interaction between electrode and condition, i.e. testing whether the age label modulates the normalized response pattern. To investigate differences in TRF pattern, the cluster-based permutation test (Maris and Oostenveld, 2007) was applied as well to identify which cluster of electrodes

drive the significant difference between the TRF patterns of older and younger participants. A significance level of $\alpha = 0.05$ was used.

3 Results

We evaluated the effect of age on acoustic and linguistic NT. First, we focused on the effect of age on NT by analysing the prediction accuracies (points 1.a. and 1.b.). Second, we investigated the temporal and spatial patterns of the neural response by exploring the TRF pattern (points 2.a. and 2.b.).

1.a. We identified whether age significantly affects the NT averaged across all electrodes. Additionally, we investigated whether the observed age effect on NT is mediated by other measures of age-related decline like hearing capacity and cognition.

1.b. We investigated the spatial pattern in NT within the a-priori selected ROIs and investigated hemispheric lateralization.

2.a. We characterized how age affects the temporal pattern of the neural response by analyzing the effect of age on the neural response latencies. Similarly, we analyzed whether the effect of age on the response latency is mediated through hearing capacity or cognition.

2.b. Lastly, we studied the associated topographies to investigate whether age affects the underlying neural source configuration responding to the different speech representations.

3.1 Behavioural results

We evaluated the correlation between age and the behavioral test results (fig. 2). We observe a positive correlation between age and PTA: with increasing age, a higher PTA value is observed (Pearson's $r = 0.478$, $p < 0.001$). Moreover, when evaluating the cognitive performance across the age span, we observed that age is positively correlated with the SCWT scores (Pearson's $r = 0.494$, $p < 0.001$) and negatively correlated with the RST outcome (Pearson's $r = -0.445$, $p < 0.001$). Both measures reflect the cognitive performance but in the opposite direction, namely, a worse cognitive performance is associated with a higher SCWT score and a lower RST outcome. Altogether the pattern of the latter two measures across age indicates that as age increases, cognitive performance declines.

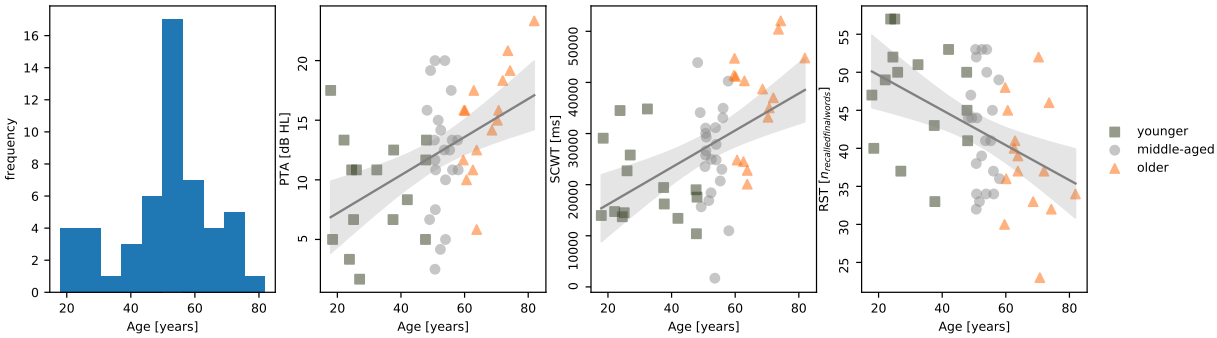

Figure 2: The distribution of age across the participants, pure tone average (PTA), Stroop color-word test (SCWT) scores and reading span test (RST) across age. The data points of 15 older adults are represented by orange triangles, while those of the 15 youngest adults are represented by green squares. All other participants, referred to as middle-aged participants, are represented by grey dots.

3.2 Effect of age on acoustic speech processing

3.2.1 Prediction accuracy

To investigate the effect of age on acoustic NT, we investigated how the prediction accuracy, averaged across time windows (30-160 ms and 160-300 ms) and over EEG electrodes, is affected by age. Acoustic NT was obtained using the forward modelling approach with the two acoustic representations: spectrogram and acoustic onsets. Using a linear model, we did not observe a significant effect of age on the prediction accuracy of the acoustic model averaged across all electrodes ($p = 0.105$). Statistics are reported in table 3.

Given that local and subtle effects cannot be detected when the prediction accuracy is averaged across all EEG electrodes, we investigated whether this age effect is observed for a priori defined sets of electrodes via a linear mixed effects model. This demonstrated that acoustic NT decreases with increasing age when NT is evaluated in the 3 ROIs, i.e., main effect of age ($p = 0.032$). The model showed no main effect of ROIs nor an interaction effect (fig. 3.B). Notably, these results differ from the previous analysis including all 64 EEG channels. This effect can be explained as only 28 of the 64 channels were included in this analysis.

Using a SEM analysis, we verified whether this age effect persists after considering the hearing capacity, i.e., after including the PTA values (more details in section Statistics under Methods). This analysis was preformed averaged across the 3 ROIs and across the two earlier time windows. Indeed, a direct effect of age on the neural tracking of acoustic representations is observed ($\beta = -0.498$, $p < 0.001$; additional details regarding this analysis are included in the supplementary material). Moreover, this analysis indicated that the effect of age is opposite to the effect of PTA. Neural tracking decreases with increasing age, while neural tracking increases with an increasing degree of hearing loss.

Opposite to previous literature, we do not find that with increasing age, acoustic NT increases. As we

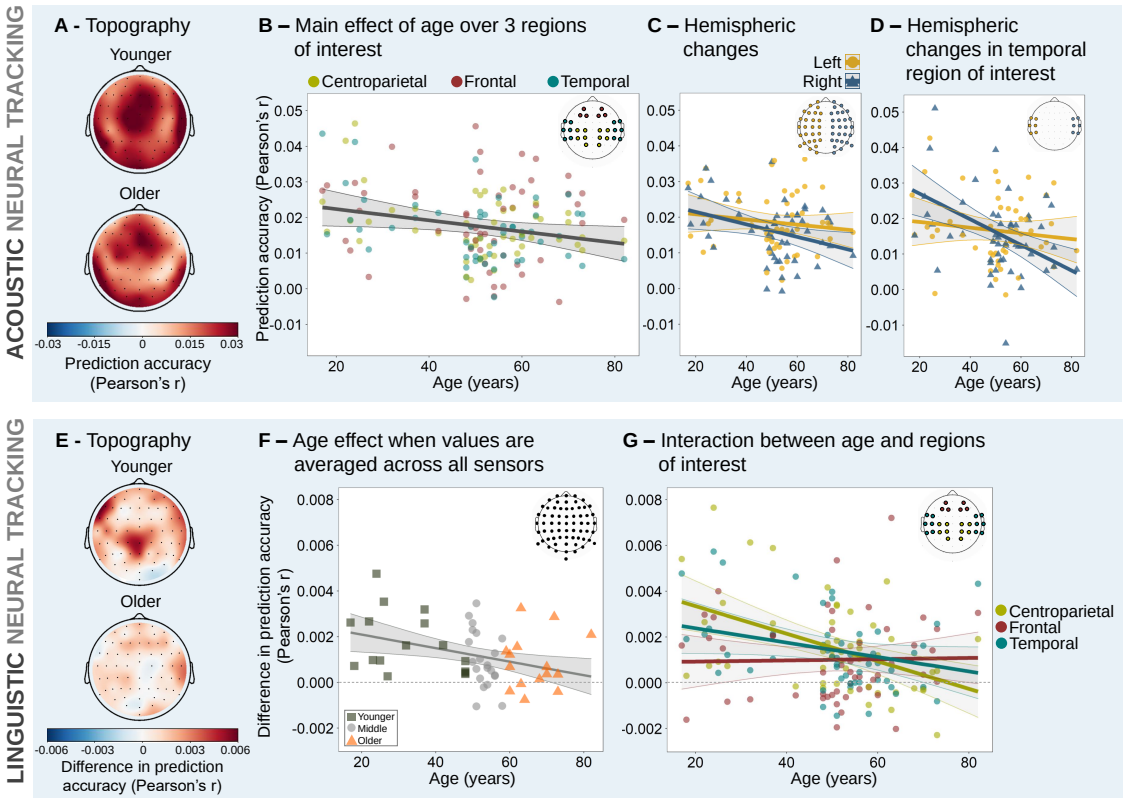

Figure 3: **Top panel A-D: Age affects acoustic neural tracking in regions of interest and interacts with hemisphere. A.** Average topography of the 15 youngest and 15 oldest participants for visualization purpose. **B.** A linear mixed effects model revealed a main effect of age over 3 ROIs, which did not differ from each other. **C.** Interaction effect between age and hemisphere when values are averaged across all left and all right hemisphere electrodes respectively. **D.** Interaction effect between age and hemisphere when values are averaged across left temporal and right temporal electrodes respectively. **Bottom panel E-G: Age affects phoneme- and word-level linguistic neural tracking. E.** Average topography of the 15 youngest and 15 oldest participants for visualization purpose. **F.** Linear age effect when values are averaged across all electrodes. **G.** A linear mixed effects model revealed an interaction between age and ROI. Post-hoc comparisons showed that the centro-parietal ROI is significantly different from the frontal one. The insets in B, C, D, F and G show the EEG electrodes included in the analysis. All statistics are reported in table 3.

are using the same data as Decruy et al. (2019), we investigated the differences in more detail. Although Decruy et al. (2019) used a different modelling approach, we aimed to replicated the finding using the forward modelling approach with similar speech representations and frequency band as used in Decruy et al. (2019). Namely, we used a frequency band of 0.5 to 8 Hz (same filter details as mentioned in the method section) and an envelope representation (calculated as the spectrogram averaged across the eight frequency bands) to allow a better comparison to the results of Decruy et al. (2019). However, similarly as above, we did not observe any effect of age on the average prediction accuracy over all electrodes using acoustic representations (model characteristics: R^2_{adj} = 0.006; F = 1.32; df = 50; p = 0.256; age-effect: β = -1.00e-04, SE = -1.75e-04, t = -1.149, p = 0.256). Moreover, when including a quadratic age term, this quadratic term does not reach significance or improve the model fit. It is noteworthy that there is a substantial methodological differences between the Decruy et al. (2019) and our results although using the same dataset. Firstly, we derived acoustic NT for a forward modeling approach, while (Decruy et al., 2019) reported acoustic NT with the backward modeling approach. In the backward modeling approach, the stimulus representations are reconstructed from the EEG data rather than the EEG is predicted based on the stimulus representations (the forward modeling approach). We used the forward modeling approach because it gives additional temporal information, i.e., how the brain responds to certain speech characteristics at specific latencies, and spatial distribution, i.e., how these brain responses are spatially distributed across the different EEG electrodes. Such spatial and temporal information is not gained using the backward modeling approach. However, the advantage of the backward model is that it optimally combines information from all EEG electrodes. Therefore, it can regress out non-stimulus-related EEG responses, resulting in more robust reconstruction accuracies. If the effect of age on NT is widespread and subtle, using across-electrode information might be essential to obtain robust results, which would explain why this effect is not observed using the forward modeling approach.

3.2.2 Hemispheric lateralization of neural tracking

Using a linear mixed effects model, a significant age and hemisphere interaction effect was found (interaction effect: p = 0.018). The age effect in the right hemisphere was significantly different from a null-model, whereas this was not the case for the left hemisphere (right: p = 0.0198; left: p = 0.324), indicating that with increasing age, participants relied less on the right hemisphere, whereas the contribution of the left hemisphere was stable across age (fig. 3C). In order to identify which spatial pattern may be responsible for this effect, we analyzed the same effect in the 3 ROIs. We found an age and hemisphere interaction effect in the temporal ROI (p = 0.049, FDR-corrected) (fig. 3D), as well as a significant main effect of age (p = 0.011, FDR-corrected), but no main effect of hemisphere (p = 0.082, FDR-corrected). Post-hoc analyses showed that in the temporal ROI, the age effect in the right hemisphere was significantly different from a null-model, whereas this was not the case for the left hemisphere

(right: p = 0.0003; left: p = 0.406). No interaction effects between age and hemisphere were found for the other 2 ROIs.

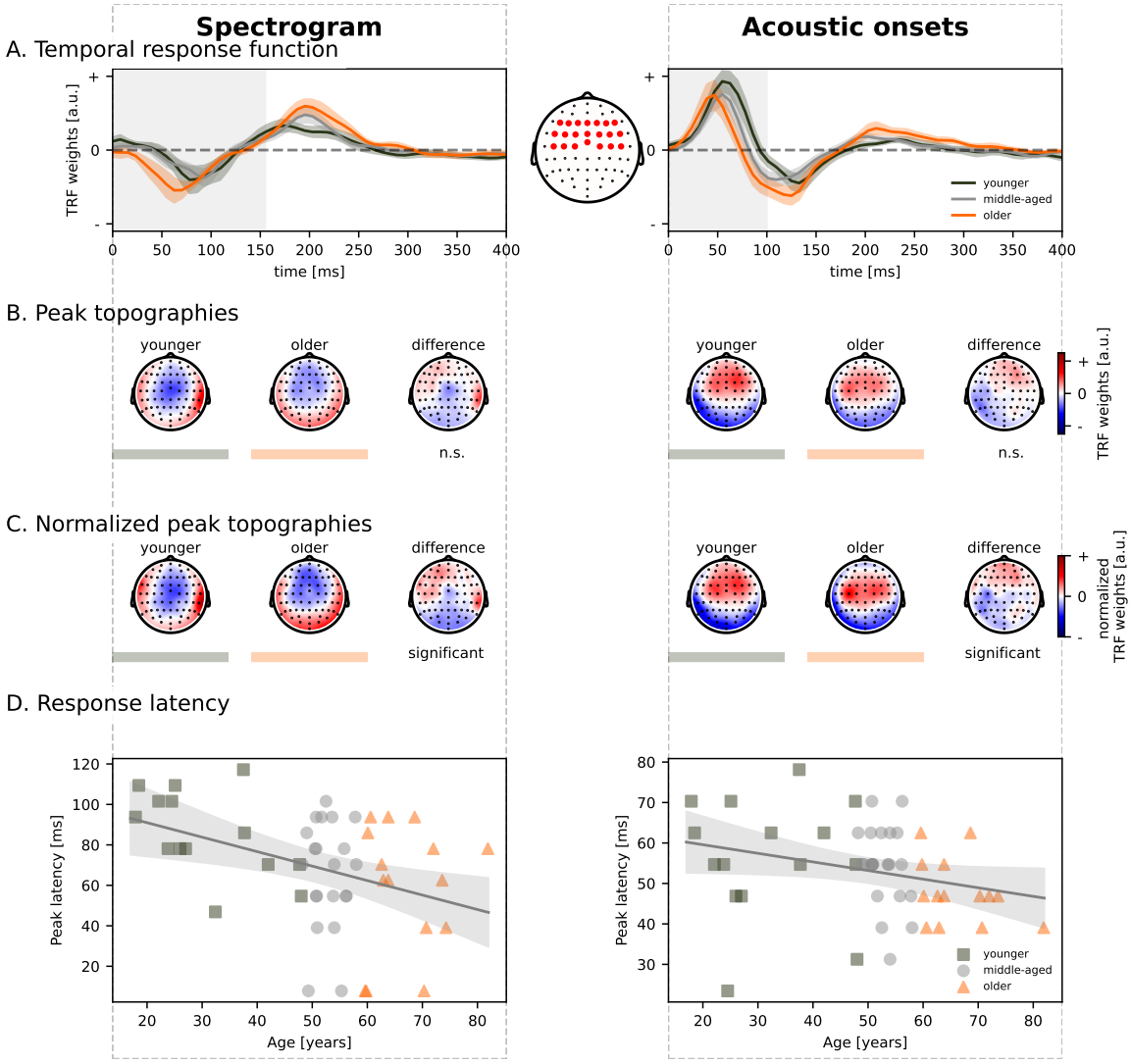

Figure 4: **Older adults show shorter latencies for the early acoustic response.** The neural responses to acoustic speech representations are visualized as a function of age for both speech representations: spectrogram (left) and acoustic onsets (right). **A.** The average temporal response functions across frontocentral electrodes (indicated on the middle inset) for younger (green), middle-aged (grey), and older (orange) participants. The shaded area indicates the time region used to find the peak topography. **B.** This panel shows the corresponding peak topographies associated with the peak found in the grey vertical panel. **C.** The corresponding normalized peak topographies associated with the peak found in the grey vertical panel. **D.** The decrease in neural response latency as a function of age. The grey lines indicate the model predictions and 95% confidence interval of how the response latency increases with age. The data points of 15 older adults are represented by orange triangles, while those of the 15 youngest adults are represented by green squares. Not all subjects showed a prominent peak; therefore, the plot does not show these data points. The insets show the corresponding peak topographies for respectively the younger and older subjects. *n.s. = not significant*

3.2.3 Neural response latency

Using a linear model, we investigated whether age impacts the latency of the neural responses to acoustic representations. As visualized in fig. 4 and confirmed by the statistical analysis, with increasing age

Table 3: Acoustic and linguistic speech processing results

Acoustic processing								
Prediction accuracy (Pearson's r)								
	Effect	**R squared**	**F-value**	**Degrees of freedom**	**Beta estimates**	**Standard error**	**t-value**	**p-value**
All electrodes	Age	0.0328†	2.733	50	-1.168e-04	7.068e-05	-1.653	0.105
ROI analysis	Whole model	0.526$^{\diamond}$						
	Age		4.8557	50				**0.032**
	ROI		1.0206	2,100				0.364
	Age*ROI		2.2033	2,100				0.116
Hemisphere analysis	Whole model	0.848$^{\diamond}$						
	Age		3.1322	50				0.083
	Hemisphere		1.4854	50				0.229
	Age*Hemisphere		5.9382	50				**0.018**
	Post-hoc: left			59	-7.27e-05	7.32e-05	-0.994	0.324$^{\triangleright}$
	Post-hoc: right			59	-1.75e-04	7.32e-05	-2.396	**0.039**$^{\triangleright}$
Temporal ROI hemisphere analysis	Whole model	0.394$^{\diamond}$						
	Age		8.0854	50				**0.011**$^{\triangleright}$
	Hemisphere		5.1722	50				0.082$^{\triangleright}$
	Age*Hemisphere		6.1720	50				**0.049**$^{\triangleright}$
	Post-hoc: left			91.4	-8.01e-05	9.59e-05	-0.835	0.406$^{\triangleright}$
	Post-hoc: right			91.4	-3.61e-04	9.59e-05	-3.761	**<0.001**$^{\triangleright}$
TRF peak latency analysis								
Spectrogram	**Age**	0.136†	8.726	48	-0.174	0.059	-2.954	**0.005**
Acoustic onsets	**Age**	0.0604†	4.216	49	-0.052	0.059	-2.954	**0.045**
Linguistic processing								
Difference in prediction accuracy (Pearson's r)								
All electrodes	**Age**	0.1061†	7.054	50	-2.962e-05	1.115e-05	-2.656	**0.011**
ROI analysis	Whole model	0.3398$^{\diamond}$						
	Age		6.4191	50				**0.014**
	ROI		6.0859	2, 100				**0.003**
	Age*ROI		4.9157	2, 100				**0.009**
	Post-hoc: centro-parietal			132	-6.03e-05	1.65e-05	-3.656	**0.001**$^{\triangleright}$
	Post-hoc: temporal			132	-3.14e-05	1.65e-05	-1.906	0.088$^{\triangleright}$
	Post-hoc: frontal			132	2.65e-06	1.65e-05	0.161	0.873$^{\triangleright}$
	Pair-wise: centro-parietal vs. frontal			100	-6.29e-05	2.01e-05	-3.132	**0.007**$^{\triangleright}$
Hemisphere analysis	Whole model	0.6342$^{\diamond}$						
	Age		6.5601	50				**0.013**
	Hemisphere		0.6776	50				0.414
	Age*Hemisphere		0.6850	50				0.412
TRF peak latency analysis								
Phoneme surprisal	**Age**	-0.0207†	0.004	48	-0.002	0.031	-0.063	0.95
Cohort entropy	**Age**	0.065†	4.413	48	0.098	0.047	2.101	**0.041**
Word surprisal	**Age**	0.101†	6.521	48	0.123	0.048	2.554	**0.014**

$\dagger$: adjusted R squared; $\diamond$: conditional R squared; $\triangleright$: FDR correction for multiple comparisons was applied; significant effects are marked in bold.

there was a decrease in latency of the early response to the spectrogram representation ($p = 0.005$) and the early response to the acoustic onsets representation ($p = 0.045$). The SEM analysis indicated that neither of these age effects were mediated through PTA.

We compared the topographies for the early response to both acoustic representations of the 15 oldest and youngest participants (visualized in fig. 4.B). For both acoustic speech representations, the topographies of the early response did not significantly differ between older and younger participants. After normalization of the topographies, i.e. removing amplitude differences between the two groups, a significant difference between the peak topographies was found between younger and older adults (normalized amplitudes are visualized in fig. 4.C; ANOVA-analysis: age-electrode interaction for the early response to the spectrogram: $F(63, 3008) = 1.77$, $p < 0.001$; age-electrode interaction for the early response to acoustic onsets: $F(126, 3072) = 2.01$, $p < 0.001$).

3.3 Effect of age on linguistic speech processing

3.3.1 Prediction accuracy

We repeated the analysis as mentioned earlier for linguistic NT. A measure of linguistic NT is calculated as the model's prediction accuracy increase when the linguistic representations are included compared to a model without these representations. This is a conservative way to determine the added value of linguistic speech representations: our tests quantify the variability *uniquely* attributable to these representations, not counting variance that is shared between acoustic and lexical segmentation representations. As a first analysis, we determined whether age affects this difference in prediction accuracy averaged across all electrodes and later integration windows (160-300 ms and 300-600 ms). Using a linear model, we observed that linguistic NT decreases with increasing age ($p = 0.011$; fig.3.F). Statistics are reported in table 3.

The linear mixed effects model used to investigate the effect of aging in 3 ROIs showed a significant main effect of age ($p = 0.014$), ROI ($p = 0.003$), as well as a significant interaction between age and ROI ($p = 0.009$). The age effect in the centro-parietal ROI was significantly different from a null-model, whereas this was not the case for the other two ROIs (centro-parietal: $p = 0.001$; temporal: $p = 0.088$; frontal: $p = 0.873$; fig.3.G). This analysis showed that with increasing age, linguistic NT decreases which is most prominent in centro-parietal areas. Pairwise comparisons revealed that the age effect in the centro-parietal ROI is significantly different from the one in the frontal ROI ($p = 0.007$) while not significantly different from the temporal ROI.

Using SEM analyses, we investigated whether the observed effects are directly due to age or whether the observed age effect in the linear models is spurious due to the correlations between age and cognition, tested via SCWT, RST, or hearing levels, tested via PTA. By investigating the linguistic NT across these two windows for both, across all electrodes and centro-parietal electrodes, this direct effect of age

persisted after including predictors for hearing capacity and cognition (more details regarding the SEM analyses in the supplementary material). These analyses suggest that the effect of decreasing linguistic NT with increasing age is likely purely due to age and not spurious due to the correlations between age with hearing capacity or cognition.

3.3.2 Hemispheric lateralization of neural tracking

Same as for the acoustic model we were interested in hemispheric lateralization changes with age. However, no significant age and hemisphere interactions were found, neither for the whole hemispheres, nor for any of the ROIs. The main effect of hemisphere was also not significant, only the main effect of age was significant ($p = 0.0135$).

3.3.3 Neural response latency to linguistic speech representations

We observed a difference in TRF between younger and older adults for the linguistic speech representations. The peaks around 250 ms and 400 ms for respectively the linguistic representations at the phoneme and word level tend to be more prominent for younger adults, while these peaks show higher variability in latency and amplitude among older adults (fig. 5). Moreover, in older adults, the peak between 100 and 200 ms seems to be enhanced. This effect is most prominent for word surprisal (fig. 6). This response around 200 ms is dominated by frontal negative activity. Surprisingly, the average TRF of the older adults does not show a significant peak around 400 ms. We attribute this to two reasons: firstly, the amplitude of the older participants in this time window is smaller, and the variability in spatial and temporal pattern of the TRF is higher across older participants. Although the amplitude is small, we could still determine a peak latency as we identified a peak based on its template topography. Within a specific time region (table 1), we searched for a topography resembling the template topography at each peak of the TRF, identified by the RMS value of the TRF as specified in section *Determination of peak latency*.

Using a linear model, we investigated whether age impacts the latency of the neural responses to linguistic speech representations. With increasing age, latency of the later response to cohort entropy increased ($p = 0.041$). A similar effect was observed for the later response to word surprisal ($p = 0.014$). These effects are visualized in fig. 5. No effects of age were found on the early responses of linguistic speech representations or on the peak latencies in the neural response to phoneme surprisal.

To investigate these age effects in more detail, we applied SEM analysis to see whether the effect of age on response latencies is still observed after including cognition and hearing capacity outcomes. The SEM analysis showed a direct age effect on the latency of the later neural response to cohort entropy. For the early response to word surprisal (125 to 255 ms) there was neither a significant direct age effect nor a total indirect age effect, although the indirect effect of age through RST reached significance (RST: $\beta =$

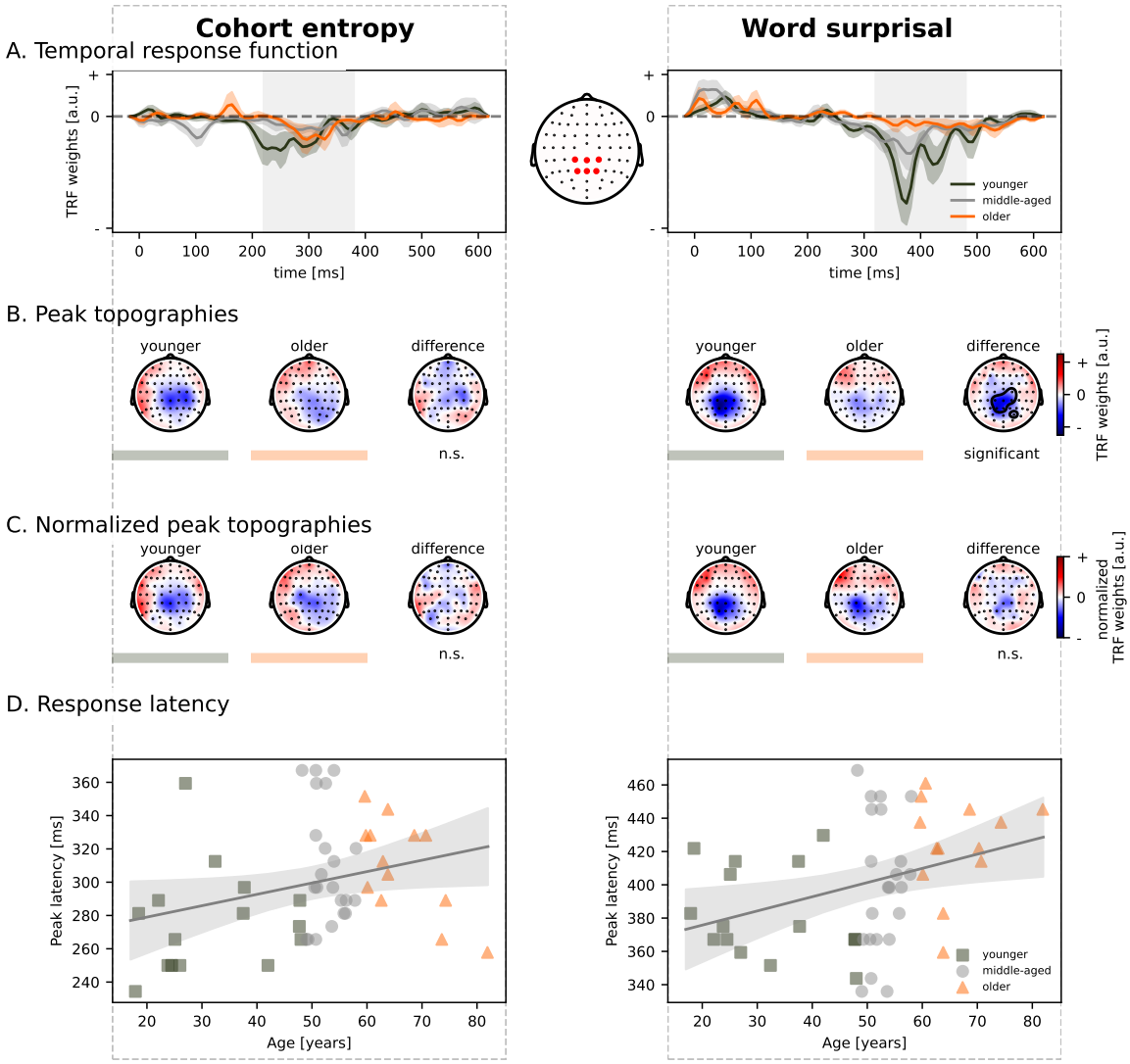

Figure 5: **Older adults show longer latencies for linguistic speech processing.** The neural responses to linguistic representations are visualized as a function of age for two linguistic speech representations: cohort entropy (left) and word surprisal (right). **A.** The average temporal response functions across central electrodes (indicated on the middle inset) for younger (green), middle-aged (grey), and older (orange) participants. **B.** The corresponding peak topographies associated with the peak found in the grey vertical panel. The contour indicates the cluster which drives the significant differences in topography between younger and older adults. **C.** The corresponding normalized peak topographies associated with the peak found in the grey vertical panel. **D.** The increase in neural response latency across age. The grey lines indicate the model predictions and 95% confidence interval of how the response latency increases with age. The data points of 15 older adults are annotated with the orange triangles, while those of the 15 youngest adults are annotated with the green squares. Not all subjects showed a prominent peak; therefore, the plot does not show these data points. *n.s. = not significant*

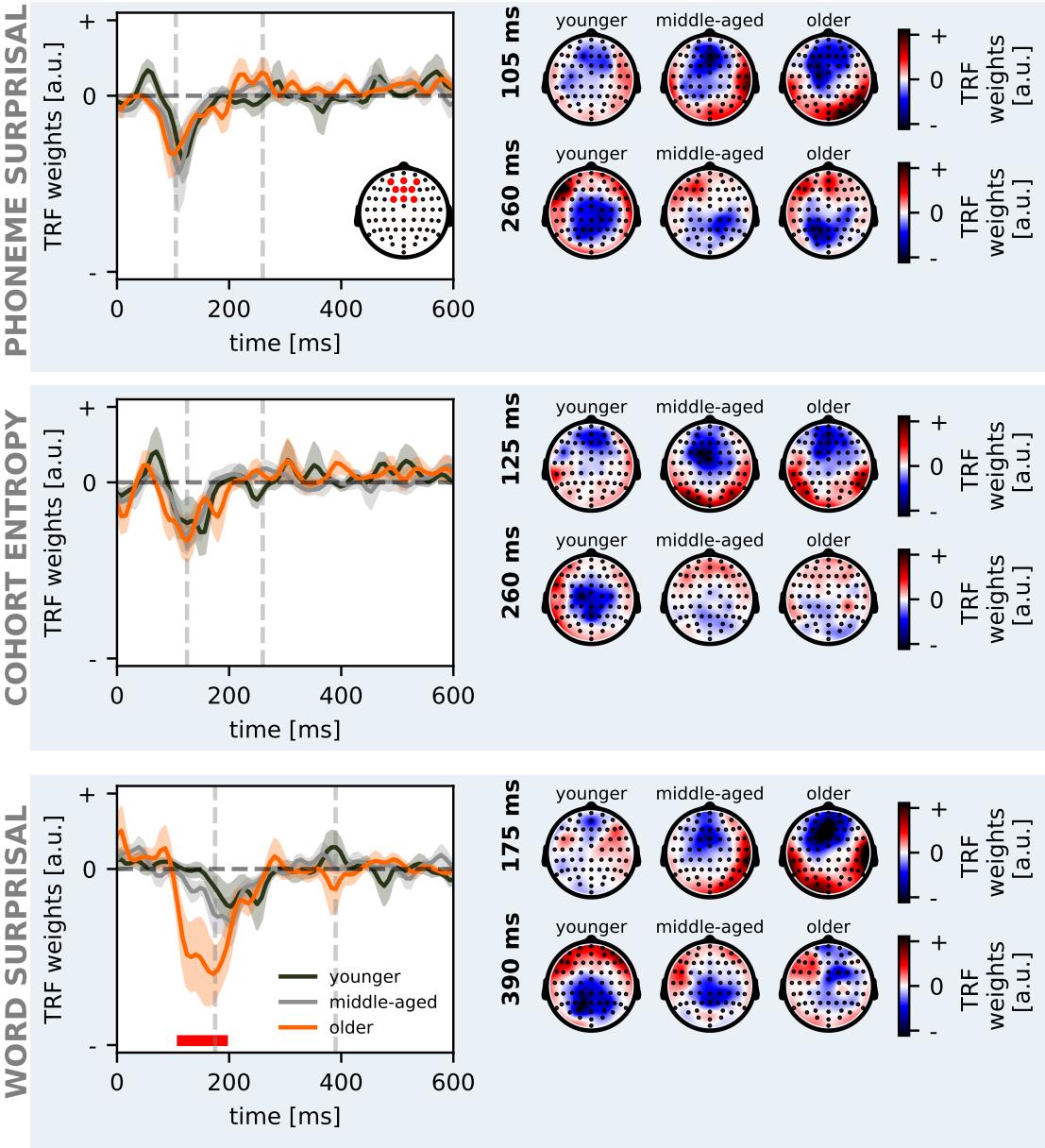

Figure 6: **Older adults show an enhanced response between 100 and 200 ms to linguistic representations.** The neural responses to linguistic representations are visualized in function of age for the three linguistic speech representations: phoneme surprisal, (top), cohort entropy (middle) and word surprisal (bottom). The left plot shows the average temporal response functions across frontocentral electrodes (indicated on the inset of the top plot) for younger (green), middle (grey), and older-aged (orange) participants. The shaded area indicated the standard error of the average TRF. The red bold line underneath the TRF indicates where the TRF is significantly different between younger and older adults. The right plot shows the corresponding peak topographies associated with the peak at the time of the dashed grey line in the left plot.

-0.057, p = 0.047). This effect indicates that as age increases, the RST performance decreases, which in turn increases the early neural response latency to word surprisal. This indicates that there is a weak effect of age mediated through the RST on the early response latency of word surprisal.For the later response to word surprisal (350 to 450 ms), no significant direct age effect was observed on the latencies of the neural responses when cognition and hearing was taken into account, with the associated indirect effect of age through SCWT almost reaching significance (indirect age effect through SCWT: $\beta = 0.051$, p = 0.055). These results indicate that the effect of age observed in the linear model is likely more due to the correlation between age and SCWT rather than a robust effect of age.

We compared the topographic responses of the 15 oldest and youngest participants (visualized as insets in fig. 5). We only observe a significant difference between younger and older adults in the topography around 400 ms to word surprisal situated centrally (cluster-based permutation test: p = 0.004). However, this significant difference disappears after the normalization of the topographies. This indicates that the significant difference observed in the permutation test is driven by an amplitude difference between younger and older adults rather than a change in topographic pattern.

4 Discussion

In order to investigate whether and how linguistic and acoustic speech processing change from early to late adulthood, we analyzed how the brain tracks respective speech representations when listening to natural speech. We modelled the unique contribution of linguistic speech representations while controlling for acoustic processing. Spatial information was investigated by analyzing a priori defined ROIs, by analyzing hemispheric lateralization and by investigating TRF topographies across age. Temporal patterns of speech integration were investigated by analyzing TRF latencies. SEM was applied to account for the influence of hearing ability and cognition on NT of natural speech.

Age affected both acoustic and linguistic speech processing. With increasing age, neural tracking generally decreased. Spatially, older adults relied less on the right hemisphere for acoustic processing, which was most prominent in temporal regions. Although no age-related change in hemispheric preference was observed for linguistic speech processing, linguistic NT decreased with increasing age, which was driven by centro-parietal regions. Exploring temporal patterns of the neural response, we found that older adults displayed earlier peaks to acoustic cues while they displayed later peaks to linguistic aspects of speech. In this normal-hearing population sample, we could not attribute these age-related differences in speech processing (both acoustic and linguistic) to an age-related decline in hearing or cognition. However, there might be some influence of cognition on the neural response latency of word-level linguistic speech processing.

4.1 Aging effects on linguistic speech processing

In this study, we show that linguistic NT decreases with increasing age. Specifically, older adults showed a lower neural tracking magnitude than younger adults in response to three linguistic speech representations, i.e., phoneme surprisal, cohort entropy and word surprisal, combined in one model. This was the case for the whole-head analysis and the centro-parietal ROI (fig. 3.E, F & G). However, no differential hemispheric lateralization with increasing age was observed for this model. The temporal analysis revealed that the TRF peak latencies of phoneme- and word-level linguistic speech representations were longer with increasing age (fig. 5). Additionally, the TRF in response to word surprisal revealed early activation over frontal electrodes in older adults only (fig. 6).

Our finding that linguistic NT decreased with increasing age deviates from previously reported results. Broderick et al. (2021) did not find a significant difference between age groups in NT of word surprisal, while Mesik et al. (2021) reported a higher neural tracking for older compared to younger adults. However, it is not straightforward to compare our results to these previously published results as neither Broderick et al. (2021) nor Mesik et al. (2021) controlled for the acoustic representation of the speech signal. Controlling for acoustics is important as the speech acoustics can explain apparent responses to higher-

level speech representations (Daube et al., 2019), leading to biased results (Gillis et al., 2021b). Based on the spatial distribution of these prediction accuracies, we observed that the effect of of decreasing linguistic speech processing across age is driven by centro-parietal areas. This is as expected as centro-parietal areas of the brain show higher NT to linguistic speech processing (Gillis et al., 2021b). Moreover, no changes in hemispheric preference were found across the lifespan.

To get additional insight into spatial and temporal aspects of how the neural responses to speech change across the lifespan, we investigated the TRF. These are in line with the N400 ERP literature, that consistently reported decreased N400 amplitudes and longer latencies in older adults (Gunter et al., 1992; Kutas and Iragui, 1998; Woodward et al., 1993; Federmeier and Kutas, 2005; Tiedt et al., 2020; Federmeier et al., 2002). However, investigating linguistic processing in natural speech has a higher ecological validity than using isolated sentences and allows to analyze lower and higher-level speech processing aspects from the same data. We show that older adults exhibit longer latencies for cohort entropy as well as for word surprisal. Although previously published studies did not consider continuous acoustic aspects of the speech, we compared the TRF in centro-parietal areas. These areas might be less biased towards speech acoustics as these affect mainly frontal areas of the brain (see electrode selection in figure 4). Broderick et al. (2021) reported similar results for word surprisal, i.e., longer latencies in older adults at the prominent negativity around 300 to 400 ms. Moreover, within the older group, the authors found a correlation between age and latency (Broderick et al., 2021). This latency shift is in agreement with the current findings. However, Mesik et al. (2021) did not observe a difference between younger and older adults in the TRF pattern to word surprisal. This could potentially be due their division into two categorical groups using an arbitrary age boundary. In studies investigating acoustic speech processing, the neural response latencies have been linked to efficiency of speech processing (Bidelman et al., 2019; Gillis et al., 2021a). Extrapolating these findings, the latency shift observed in linguistic speech processing might indicate that with increasing age, linguistic neural processing becomes more complex, i.e., more neural processing time is required to integrate a phoneme or word into the story context. In essence, same as for the magnitude of linguistic NT, age also impacts the latency of linguistic speech processing.

One can ask which aspects of healthy aging drive these differences in linguistic NT and latency shifts: cognition, age, or a complex interaction between both. The effect of age on linguistic NT, i.e., the prediction accuracy, was neither mediated by cognition nor hearing capacity. However, the latency shift observed for word surprisal was driven by cognition rather than age. At the word-level, a similar effect was reported by Federmeier and Kutas (2005): longer latencies for the N400 were observed for older adults with a lower RST score. Although the latency effect in that study was driven by another cognitive predictor, i.e. RST instead of SCWT score, altogether, this might indicate a link between decreased cognition and longer latencies. Surprisingly, these results show a discrepancy between these two SEM results, i.e., the mediation through age is different for prediction accuracies than for neural response

latencies. This indicates that the sensitivity of the used measures of hearing capacity and cognition on NT and the neural response latency differed. With increasing age, we observed decreased linguistic NT, which was not mediated through hearing capacity nor cognition. Linguistic NT was evaluated across the two later windows, and therefore it combines the effect at phoneme and word level. Furthermore, a change in the prediction accuracy is not necessarily related to a shift in the neural response latency. Therefore, the changes in the TRF pattern might give additional insights. Indeed, the SEM analysis indicated that only the latency shift to word surprisal was driven by cognition while this was not the case for cohort entropy. These results suggest that specific aspects of language processing at the phoneme level might be affected by an age effect that is not explained by cognitive decline. For specific aspects of language processing at the word level, slower responses might be due to lower scores on cognition rather than an age-decrease in predictive coding of words.

These changes in linguistic NT and in neural response latency might be related to the changes in cognition and brain structure that occur during the healthy aging process. Age-related structural changes can cause reduced coordination between recruited brain regions to process speech (Peelle et al., 2010). White matter tracts are the brain's tool kit to communication and coordination within and between regions. However, decreased white matter directionality (Salthouse, 2011; Voineskos et al., 2012), increased white matter lesion volume (Bennett and Madden, 2014; Tubi et al., 2020) and concomitant decreased processing speed (Salthouse, 1996; Salami et al., 2012; Salthouse, 2011; Bennett and Madden, 2014) are side effects of healthy aging and have been linked to age-related cognitive decline and verbal processing in older adults (Hoffman et al., 2017). Therefore, there is an intrinsic link between the age-related structural changes in white matter, functional changes apparent in delays of neural processes and behavioural measures of cognition. Our results point to the same conclusion: structural changes are indicated by the results of the topographic analyses, functional changes are shown by the differences in neural response latencies and the results of the SEM analysis creates the link with cognition. Our results supporte this hypothesis that there is an intrinsic relation between changes in cognition and brain structure. The SEM analysis indicated that behaviourally-measured aspects of cognition are related to the neural response latency to word surprisal, i.e., suggesting lower processing speed for adults with a higher cognitive decline. This lower processing speed is a trace of functional changes in the brain. At the same time, the topographic analysis showed that the topographic response to the later response of word surprisal differed between younger and older adults suggesting different underlying neural activity, which might be the result of structural changes. However, as we did not conduct source localization and did not have access to structural MRI images or white matter metrics, limited insight can be gained into the structural changes of the brain. In future studies, it might therefore be interesting to assess processing speed and/or take into account white matter metrics.

Compensating for the age-related decline in brain structure by using different, more cost-efficient neural

strategies can help elderly to maintain a viable performance level (Reuter-Lorenz and Cappell, 2008; Diaz et al., 2016; Cabeza et al., 2018). The presently observed decrease in linguistic NT with advancing age may be due to a change in neural strategy. In fact, the age-related change in the N400 ERP has been explained by use of a different strategy for extracting sentence context information (Federmeier et al., 2003). To facilitate speech, older adults may rely more on a mechanism called plausibility than on generating predictions (Federmeier et al., 2002). Prediction is referred to as the access or activation of word meaning, while plausibility refers to the integration of word meaning in its linguistic context, i.e. judging how plausible a word is within a sentence (Federmeier et al., 2002; Nieuwland et al., 2020). The linguistic speech representations used here may not capture both mechanisms to the same extent, which could have an influence on our results. Future studies could investigate age-related processing differences between these different mechanisms used in linguistic speech processing in response to continuous speech. The present spatial analysis of linguistic NT as well as the N400 peak topography (fig. 5.B) show that the magnitude decrease with advancing age is specific to the centro-parietal ROI, which is the same region in which the N400 ERP typically displays (Kutas and Federmeier, 2011; Nieuwland et al., 2020). The N400 topography to word surprisal showed a significantly different peak topography in younger and older adults (fig. 5.B). This suggests that the magnitude of activation of the underlying neural source configuration changes with age. One explanation might be that older adults rely less on prediction than plausibility which might be less well captured by the chosen linguistic representations. Interestingly, older adults, but not younger adults, displayed a peak between 100-200ms, which is modulated by how surprising a word is in its sentence context (fig. 6). This higher neural activation showed a prominent negativity in frontal areas. We can only speculate which process causes this. One explanation is age-related neurophysiological changes in the brain (Peelle et al., 2010): older adults have a reduced ability to recruit regions in the left ventral inferior frontal gyrus. Therefore, older adults recruit additional frontal regions related to working memory support to compensate for this loss. Although we did not perform source localization, the early neural activity to word surprisal in older adults may resemble the additional recruitment of frontal areas related to working memory. The SEM analysis indicated a link between the RST score and early word surprisal peak latency (for details see supplementary material section 6.2.2 on page 47 and fig. S.2 on page 49).

A change in the areas recruited to process linguistic aspects of speech could also have an impact on the hemispheric lateralization. In fact, Diaz et al. (2016) suggested that older adults additionally recruit the right inferior frontal gyrus (Wierenga et al., 2008; Grossman et al., 2002; Wingfield and Grossman, 2006) for processes relying on the left-lateralized dorsal stream in younger adults (Hickok and Poeppel, 2007), e.g., linguistic speech processing among others (Kandylaki et al., 2016). Therefore, we expected more bilateral linguistic speech processing in older than in younger adults. However, we did not find hemispheric lateralization changes across age for linguistic speech processing, meaning that younger

as well as older adults had symmetrical results. Even though this is not in line with the theoretical frameworks in younger adults proposed by Hickok and Poeppel (2007) and Wlotko et al. (2010), this result is not necessarily surprising given that the N400 measured with EEG has consistently been found over centro-parietal electrodes (Kutas and Federmeier, 2011; Broderick et al., 2021). In older adults, the fact that neither hemisphere is specialized in linguistic speech processes aligns with the literature, suggesting that the elderly rely on bilateral neural sources for this function (Wlotko et al., 2010; Diaz et al., 2016).

A caveat of the current study is the limited amount of data. All participants listened to the same 12-minute-long story. Twelve minutes is not much to train all the used speech representations. Previous literature investigating linguistic representations relied 30 minutes or more data. Nevertheless, a significant trend across age is observed: the unique contribution of linguistic representations decreases with increasing age. Either of two reasons might explain this: (1) It might be that the current linguistic representations do not model the underlying neural responses to linguistic speech processing in older adults, referring to predictability versus plausibility, as raised by Nieuwland et al. (2020). (2) The proportion of EEG signals related to linguistic speech processing in the measured EEG activity might be decreased in older adults. Such a decrease might be due to a decrease in the magnitude of the neural sources related to linguistic speech processing or due to an increase in brain noise (Voytek et al., 2015). Since the contribution of linguistic representations decreases with increasing age, the interpretation of the TRF should be done with care. The contribution of linguistic representations in older adults is nearly zero, indicating that the TRF has little explanatory power. Therefore the observed effects might not be robust. Nevertheless, the observed trends in this TRF pattern align with previous studies.

In essence, the current results provide evidence for a decline in linguistic speech processing with advancing age and this is likely caused by compensatory mechanisms, i.e., usage of different functional strategies and recruitment of different neural structures related to age-related structural changes in the brain. Moreover, this compensation mechanism might be modulated by the listeners' cognitive functioning, at least at word-level processing.

4.2 Aging effects on acoustic speech processing

This study shows that acoustic NT is not affected by age when considering all electrodes. However, when averaging over frontal, temporal and centro-parietal electrodes, acoustic NT decreases with advancing age (fig. 3.B). Hitherto, speech envelope tracking studies that have investigated aging trends, have reported results opposing the current ones, namely increased acoustic NT with increasing age (Presacco et al., 2016; Brodbeck et al., 2018b; Decruy et al., 2019; Karunathilake et al., 2020). This is however likely due to methodological differences, e.g., a different modelling approach (backward versus forward modelling), different acoustic speech representations (envelope versus spectrogram and acoustic edges) or a different

composition of the studied sample (age categories versus age as continuous variable).

The temporal pattern analysis unveiled shorter TRF latencies with increasing age for the early peak of the acoustic speech representations, i.e. spectrogram and acoustic onsets (fig. 4.A & C). Younger adults thus showed a later acoustic response than older adults. Based on previous literature using simple sounds as clicks, tones, or syllables, we expected to find an age difference in the later acoustic response (N1 and P2, around 100-200ms) rather than the early acoustic response (before 100ms). In general, we hypothesized longer latencies with advancing age, but the current results demonstrate the opposite. One possible explanation might be the task demand: listening to speech is an automatic process that humans are highly trained at, whereas processing artificially created clicks, tones or syllables is less familiar and might thus influence the neural response differently. Moreover, neural speech processing is modulated by attention effects. Older adults may be more involved with the task than younger adults, resulting in shorter neural response latencies to acoustic speech processing. Using word pairs, Roque et al. (2019) also observed a decrease in early acoustic response latency in older adults.

The normalized topographies of these early acoustic responses significantly differed between the 15 youngest and 15 oldest adults (fig. 4B). This finding supports that the neural activity underlying this early acoustic response is inherently different between younger and older adults. A topography reflects a complex interplay between the neural activity of different sources. Therefore, a significant difference in normalized topography suggests that either the location of the neural activity differed and/or the relative magnitude of activation between the different neural sources changed. Although Brodbeck et al. (2018b) did not report a latency difference, their results showed that older adults recruited additional regions at a short latency, around 30 ms, relative to the slow amplitude modulations of the speech envelope. They suggested that higher envelope tracking with increasing age might not necessarily imply enhanced encoding of acoustic representations but might indicate increased neural activity in the left temporal lobe. Likewise, Farahani et al. (2020) found enhanced activity in older adults in the left hemisphere for processing simple, non-speech, acoustic stimuli presented at the syllable rate. Additionally, we showed evidence that older adults recruit different neural sources coming from the analysis investigating hemispheric lateralization. Herein, we reported that with increasing age, participants rely less on the right hemisphere for acoustic speech processing, which is most prominent in temporal regions. These results also align with the findings of Brodbeck et al. (2018b) and Farahani et al. (2020) in the sense that older adults recruit additional or different neural sources in the left hemisphere for acoustic processing of speech. Interestingly, we found a left-hemispheric lateralization with advancing age for acoustic NT, but not for linguistic NT, which indicates that the additional activity in the left temporal cortex in older adults as reported by Brodbeck et al. (2018b) does not seem to be linked to linguistic speech processing mechanisms, but rather to acoustic processing. Altogether, this evidence suggests that older adults recruit different neural sources to process speech sounds compared to younger adults.

Our temporal and spatial findings suggest that the neural resources involved in acoustic speech processing change across the lifespan. This points towards an age-related compensation mechanism. In fact, structural magneto-resonance imaging (MRI) research shows that gray matter volume reductions in areas for auditory processing have been found in the temporal lobe, specifically in Heschl's gyrus, planum temporale and superior temporal gyrus (Harris et al., 2009; Tremblay et al., 2013; Meunier et al., 2014; Profant et al., 2014; Cardin, 2016). Furthermore, reduced directionality of white matter fibers in older adults has been found in the acoustic radiation, Heschl's gyrus and superior temporal gyrus (Lutz et al., 2007). This evidence together with the findings of the current study suggest that older adults might draw on different neural resources to compensate for such structural changes.

5 Conclusion

This study illustrates that both acoustic and linguistic speech processing mechanisms change across the adult lifespan, which is likely underpinned by recruitment of different neural sources. For the first time, the unique contribution of NT of linguistic speech processing was investigated in older adults. We demonstrated that linguistic speech processing declines with advancing age. Moreover, as age increased, the neural response latency to certain aspects of linguistic speech processing increased. These age-related changes were mainly apparent over centro-parietal electrodes. Acoustic NT decreased as well with increasing age, but in contrast to linguistic processing older subjects showed shorter latencies for the early acoustic response to speech. Different neural resources were involved in acoustic speech processing in older than in younger adults, e.g., less involvement of the right hemisphere with advancing age.

Furthermore, we observed that linguistic speech processing in older adults might be affected by a complex interplay between age-related structural changes and functional changes which might be modulated by cognition at the word-level. In future studies analyzing neural speech tracking across the adult lifespan, it would be interesting to explore this complex interplay by taking into account additional behavioral correlates of cognition and by using source localization and functional connectivity methods in experiments using natural, continuous speech.

Our results indicate that linguistic neural tracking decreases across the lifespan. This study shows the need to investigate further how linguistic mechanisms change during aging to find a robust measure of speech comprehension in older adults. Such a measure of linguistic processing would significantly impact the diagnosis and speech therapy of people with language and communication disorders, such as aphasia, which predominantly affect older adults.

Data Availability Statement

After completion of the review process the preprocessed data, i.e. the neural tracking values, neural response latencies, etc., will be made available. The code for the template matching is available via the GitHub repository (https://github.com/MarliesG/TemplateMatching).

Funding Sources

The presented study received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Tom Francart; grant agreement No. 637424). Research of Marlies Gillis (PhD grant: SB 1SA0620N) was funded by the Research Foundation Flanders (FWO). Research of Jill Kries was supported by the Luxembourg National Research Fund (FNR) (AFR-PhD project reference 13513810). Furthermore, this study was financially supported by the FWO, grant

No. G0D8520N.

Disclosures

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

Acknowledgements

The authors would like to thank Dr. Toivo Glatz for methodological advice, Prof. Dr. Astrid Van Wieringen for advice on the cognitive part of the data and Klara Schevenels for providing valuable comments on the introduction and discussion. We would also like to thank Dr. Lien Decruy and Dr. Jonas Vanthornhout for collecting the dataset used in this study. They were assisted in data collection by Elien Van den Borre, Melissa Schoubben, Dr. Annelies Devesse, and Dr. Sam Denys. A thank you also goes to all participants.

Author contributions

TF, JK, MG and MV conceived and designed research; JK and MG analyzed data; JK, MG, TF and MV interpreted results of experiments; JK and MG prepared figures; JK and MG drafted manuscript; MV, TF, MG and JK edited and revised manuscript; MV, TF, MG and JK approved final version of manuscript.

References

Alday, P. M. (2019). M/EEG analysis of naturalistic stories: a review from speech to language processing. *Language, Cognition and Neuroscience*, 34(4):457–473.

Barton, K. (2022). *Multi-Model Inference*. R package version 1.46.0.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 (1995), Publi. *Journal of the Royal Statistical Society*, 57(1):289–300.

Bennett, I. J. and Madden, D. J. (2014). Disconnected aging: Cerebral white matter integrity and age-related differences in cognition. *Neuroscience*, 276:187–205.

Bidelman, G. M., Price, C. N., Shen, D., Arnott, S. R., and Alain, C. (2019). Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. *Hearing research*, 382:107795.

Brennan, J. R. and Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLoS ONE*, 14(1):1–18.

Brodbeck, C. (2020). Eelbrain 0.32. `http://doi.org/10.5281/zenodo.3923991`.

Brodbeck, C., Bhattasali, S., Heredia, A. C., Resnik, P., Simon, J. Z., and Lau, E. (2021). Parallel processing in speech perception: Local and global representations of linguistic context. *bioRxiv*, page 2021.07.03.450698.

Brodbeck, C., Hong, L. E., and Simon, J. Z. (2018a). Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology*, 28(24):3976–3983.

Brodbeck, C., Presacco, A., Anderson, S., and Simon, J. Z. (2018b). Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acustica united with Acustica*, 104(5):774–777.

Brodbeck, C. and Simon, J. Z. (2020). Continuous speech processing. *Current Opinion in Physiology*, 18:25–31.

Broderick, M. P., Di Liberto, G. P., Anderson, A. J., Rofes, A., and Lalor, E. C. (2021). Dissociable electrophysiological measures of natural language processing reveal differences in comprehension strategy in healthy ageing. *Scientific reports*.

Burke, D. M. and Shafto, M. A. (2008). Language and aging. In *The handbook of aging and cognition, 3rd ed.*, pages 373–443. Psychology Press, New York, NY, US.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., and Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, 19(11):701–710.

Cardin, V. (2016). Effects of aging and adult-onset hearing loss on cortical auditory regions. *Frontiers in Neuroscience*, 10:199.

Daube, C., Ince, R. A., and Gross, J. (2019). Simple acoustic features can explain phoneme-based predictions of cortical responses to speech. *Current Biology*, 29(12):1924–1937.

David, S. V., Mesgarani, N., and Shamma, S. A. (2007). Estimating sparse spectro-temporal receptive fields with natural stimuli. *Network: Computation in neural systems*, 18(3):191–212.

Decruy, L., Vanthornhout, J., and Francart, T. (2019). Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *Journal of neurophysiology*, 122(2):601–615.

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

Diaz, M. T., Rizio, A. A., and Zhuang, J. (2016). The Neural Language Systems That Support Healthy Aging: Integrating Function, Structure, and Behavior. *Language and Linguistics Compass*, 10(7):314–334.

Duchateau, J., Kong, Y. O., Cleuren, L., Latacz, L., Roelens, J., Samir, A., Demuynck, K., Ghesquière, P., Verhelst, W., and Van hamme, H. (2009). Developing a reading tutor: Design and evaluation of dedicated speech recognition and synthesis modules. *Speech Communication*, 51(10):985–994.

Farahani, E. D., Wouters, J., and van Wieringen, A. (2020). Neural Generators Underlying Temporal Envelope Processing Show Altered Responses and Hemispheric Asymmetry Across Age. *Frontiers in Aging Neuroscience*, 12(December):1–19.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4):491–505.

Federmeier, K. D. and Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*.

Federmeier, K. D., McLennan, D. B., de Ochoa, E., and Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2):133–146.

Federmeier, K. D., Schwartz, T. J., Van Petten, C., and Kutas, M. (2003). Sounds, Words, Sentences: Age-Related Changes Across Levels of Language Processing. *Psychology and Aging*, 18(4):858–872.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

Fuglsang, S. A., Märcher-Rørsted, J., Dau, T., and Hjortkjær, J. (2020). Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. *Journal of Neuroscience*, 40(12):2562–2572.

Gillis, M., Decruy, L., Vanthornhout, J., and Francart, T. (2021a). Hearing loss is associated with delayed neural responses to continuous speech. *bioRxiv*.

Gillis, M., Vanthornhout, J., Simon, J. Z., Francart, T., and Brodbeck, C. (2021b). Neural markers of speech comprehension: measuring eeg tracking of linguistic speech representations, controlling the speech acoustics. *bioRxiv*.

Goossens, T., Vercammen, C., Wouters, J., and van Wieringen, A. (2016). Aging affects neural synchronization to speech-related acoustic modulations. *Frontiers in Aging Neuroscience*, 8(JUN).

Grossman, M., Cooke, A., DeVita, C., Alsop, D., Detre, J., Chen, W., and Gee, J. (2002). Age-related changes in working memory during sentence comprehension: an fMRI study. *NeuroImage*, 15(2):302–317.

Gunter, T. C., Jackson, J. L., and Mulder, G. (1992). An Electrophysiological Study of Semantic Processing in Young and Middle-Aged Academics. *Psychophysiology*, 29(1):38–54.

Gwilliams, L. and Davis, M. H. (2022). Extracting language content from speech sounds: the information theoretic approach. In *Speech Perception*, pages 113–139. Springer.

Hamilton, L. S., Oganian, Y., Hall, J., and Chang, E. F. (2021). Parallel and distributed encoding of speech across human auditory cortex. *Cell*, 184:4626–4639.

Hammes, J. G. W. (1978). *De Stroop kleur-woord test*. Harcourt Test Publ.

Harris, K. C. (2020). *The Aging Auditory System: Electrophysiology*, pages 117–141. Springer International Publishing, Cham.

Harris, K. C., Dubno, J. R., Keren, N. I., Ahlstrom, J. B., and Eckert, M. A. (2009). Speech recognition in younger and older adults: a dependency on low-level auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(19):6078–6087.

Heeris, J. (2014). Gammatone filterbank toolkit 1.0. `https://github.com/detly/gammatone`.

Hickok, G. and Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(May):393–402.

Hoffman, P., Cox, S. R., Dykiert, D., Muñoz Maniega, S., Valdés Hernández, M. C., Bastin, M. E., Wardlaw, J. M., and Deary, I. J. (2017). Brain grey and white matter predictors of verbal ability traits in older age: The Lothian Birth Cohort 1936. *NeuroImage*, 156(November 2016):394–402.

Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.

Kandylaki, K. D., Nagels, A., Tune, S., Kircher, T., Wiese, R., Schlesewsky, M., and Bornkessel-Schlesewsky, I. (2016). Predicting "when" in discourse engages the human dorsal auditory stream: An fMRI study using naturalistic stories. *Journal of Neuroscience*, 36(48):12180–12191.

Karunathilake, I. D., Dunlap, J. L., Perera, J., Presacco, A., Decruy, L., Anderson, S., Kuchinsky, S. E., and Simon, J. Z. (2020). Effects of aging on the cortical representation of continuous speech. *AROposter*.

Kuperberg, G. R. and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1).

Kutas, M. and Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1):621–647.

Kutas, M. and Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography & Clinical Neurophysiology: Evoked Potentials*, 108(5):456–471.

Lenth, R. V., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., and Singmann, H. (2022). *Estimated Marginal Means, aka Least-Squares Means.* R package version 1.7.3.

Lutz, J., Hemminger, F., Stahl, R., Dietrich, O., Hempel, M., Reiser, M., and Jäger, L. (2007). Evidence of subcortical and cortical aging of the acoustic pathway: a diffusion tensor imaging (DTI) study. *Academic radiology*, 14(6):692–700.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, 164(1):177–190.

Martin, B. A., Tremblay, K. L., and Korczak, P. (2008). Speech Evoked Potentials : From the Laboratory to the Clinic. *Ear and Hearing*, 29:285–313.

McCarthy, G. and Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 62(3):203–208.

McCullagh, J. and Shinn, J. B. (2013). Auditory cortical processing in noise in younger and older adults. *Hearing, Balance and Communication*, 11(4):182–190.

Mesik, J., Ray, L., and Wojtczak, M. (2021). Effects of Age on Cortical Tracking of Word-Level Features of Continuous Competing Speech. *Frontiers in Neuroscience*, 15(April):1–21.

Meunier, D., Stamatakis, E. A., and Tyler, L. K. (2014). Age-related functional reorganization, structural changes, and preserved cognition. *Neurobiology of Aging*, 35(1):42–54.

Nasreddine, Z. (2004). *Montreal cognitive assessment (MoCA)*. École des sciences de la réadaptation, Sciences de la santé, Université d'Ottawa.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., Huettig, F., Husband, E. M., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., Rueschemeyer, S. A., Segaert, K., Tuomainen, J., and Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a largescale study using brain potentials. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1791).

Peelle, J. E., Gross, J., and Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, 23(6):1378–1387.

Peelle, J. E., Troiani, V., Wingfield, A., and Grossman, M. (2010). Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cerebral Cortex*, 20(4):773–782.

Presacco, A., Simon, J. Z., and Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of neurophysiology*, 116(5):2346–2355.

Prinsloo, K. D. and Lalor, E. C. (2020). General auditory and speech-specific contributions to cortical envelope tracking revealed using auditory chimeras. *bioRxiv*.

Profant, O., Škoch, A., Balogová, Z., Tintěra, J., Hlinka, J., and Syka, J. (2014). Diffusion tensor imaging and MR morphometry of the central auditory pathway and auditory cortex in aging. *Neuroscience*, 260:87–97.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reuter-Lorenz, P. A. and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current Directions in Psychological Science*, 17(3):177–182.

Roque, L., Karawani, H., Gordon-Salant, S., and Anderson, S. (2019). Effects of age, cognition, and neural encoding on the perception of temporal speech cues. *Frontiers in neuroscience*, 13:749.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2):1–36.

RStudio Team (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA.

Salami, A., Eriksson, J., Nilsson, L. G., and Nyberg, L. (2012). Age-related white matter microstructural differences partly mediate age-related decline in processing speed but not cognition. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1822(3):408–415.

Salthouse, T. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychol Bull*, 137(5):753–784.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological review*, 103(3):403–428.

Shafto, M. A. and Tyler, L. K. (2014). Decline and Preservation. *Science*, 346(6209):583–588.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., and Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, 138:1–52.

Slaney, M. (1998). Auditory toolbox. *Interval Research Corporation, Tech. Rep*, 10(1998).

Somers, B., Francart, T., and Bertrand, A. (2018). A generic eeg artifact removal algorithm based on the multi-channel wiener filter. *Journal of neural engineering*, 15(3):036007.

Spreng, R. N. and Turner, G. R. (2019). The Shifting Architecture of Cognition and Brain Function in Older Adulthood. *Perspectives on Psychological Science*, 14(4):523–542.

Tiedt, H. O., Ehlen, F., and Klostermann, F. (2020). Age-related dissociation of n400 effect and lexical priming. *Scientific Reports*, 10(1):1–10.

Tremblay, K. L., Piskosz, M., and Souza, P. (2002). Aging alters the neural representation of speech cues. *NeuroReport*, 13(15):1865–1870.

Tremblay, K. L., Piskosz, M., and Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology*, 114(7):1332–1343.

Tremblay, P., Dick, A. S., and Small, S. L. (2013). Functional and structural aging of the speech sensorimotor neural system: functional magnetic resonance imaging evidence. *Neurobiology of aging*, 34(8):1935–1951.

Tubi, M. A., Feingold, F. W., Kothapalli, D., Hare, E. T., King, K. S., Thompson, P. M., and Braskie, M. N. (2020). White matter hyperintensities and their relationship to cognition: Effects of segmentation algorithm. *NeuroImage*, 206(November 2019):1–11.

van den Noort, M., Bosch, P., Haverkort, M., and Hugdahl, K. (2008). A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment*, 24(1):35–42.

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., and Francart, T. (2018). Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. *JARO - Journal of the Association for Research in Otolaryngology*, 19(2):181–191.

Vercammen, C., van Wieringen, A., Wouters, J., and Francart, T. (2017). Desynchronisation of auditory steady-state responses related to changes in interaural phase differences: an objective measure of binaural hearing. *International journal of audiology*, 56(7):464–471.

Voineskos, A. N., Rajji, T. K., Lobaugh, N. J., Miranda, D., Shenton, M. E., Kennedy, J. L., Pollock, B. G., and Mulsant, B. H. (2012). Age-related decline in white matter tract integrity and cognitive performance: A DTI tractography and structural equation modeling study. *Neurobiology of Aging*, 33(1):21–34.

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., and Gazzaley, A. (2015). Age-related changes in 1/f neural electrophysiological noise. *Journal of Neuroscience*, 35(38):13257–13265.

Weissbart, H., Kandylaki, K. D., and Reichenbach, T. (2019). Cortical tracking of surprisal during continuous speech comprehension. *Journal of Cognitive Neuroscience*, 32(1):155–166.

Wierenga, C. E., Benjamin, M., Gopinath, K., Perlstein, W. M., Leonard, C. M., Rothi, L. J., Conway, T., Cato, M. A., Briggs, R., and Crosson, B. (2008). Age-related changes in word retrieval: Role of bilateral frontal and subcortical networks. *Neurobiology of Aging*, 29(3):436–451.

Wingfield, A. and Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology*, 96(6):2830–2839.

Wlotko, E. W., Lee, C. L., and Federmeier, K. D. (2010). Language of the Aging Brain: Event-Related Potential Studies of Comprehension in Older Adults. *Linguistics and Language Compass*, 4(8):623–638.

Woodward, S. H., Ford, J. M., and Hammett, S. C. (1993). N4 to spoken sentences in young and older subjects. *Electroencephalography and clinical neurophysiology*, 87 5:306–320.

6 Supplementary Material

6.1 Structural Equation Modelling - Acoustics

6.1.1 Neural Tracking

Using these standardized estimates, we applied SEM to assess the effect of PTA and age on the prediction accuracy averaged across the three ROIs and across the two time windows: 30-160 ms and 160-300 ms. We observed a marginal indirect effect of age through the hearing capacity (represented by PTA) on the average NT in centro-parietal and temporal electrode selections (model specifications: RMSEA = 0, CFI = 1, TLI = 1; saturated model; indirect effect; β = 0.0.158, SE = 0.076, z-value = 2.006, p = 0.045). The PTA is affected by age (β = 0.478, SE = 0.122, z-value = 3.921, p $<$ 0.001), and PTA affects the NT (β = 0.313, SE = 0.138, z-value = 2.269, p = 0.023). We observed both a direct effect of age on acoustic NT: as age increases, acoustic NT decreases (direct effect; β = -0.498, SE = 0.141, z-value = -3.527, p $<$ 0.001). A schematic overview of the SEM results is visualized in the left panel of fig. S.1.

Therefore, even though the degree of hearing loss is taken into account, a direct effect of age on the neural tracking centro-parietal and temporal electrodes is still observed. Interestingly, the indirect effect of age has an opposite sign than the direct age effect. As age increases, the degree of hearing loss increases reflected by higher PTA values which increases the neural tracking. The latter assumption converges with studies investigating the impact of hearing loss: hearing-impaired listeners show an enhanced neural tracking of speech compared to normal-hearing listeners (Decruy et al., 2020; Fuglsang et al., 2020).

6.1.2 Neural Response Latency

We investigated the SEM for each peak latency for each speech representation separately. As age is highly correlated with the degree of hearing loss, SEM allows us to model the effect of age directly onto the response latency as well as the effect of age mediated through the PTA values. The effect of age was not mediated through PTA for any peak latency. More details regarding these analyses are given in the following paragraphs (schematic overview of the SEM results are visualized in fig. S.1).

For the response latency of the first peak of the spectrogram (30 to 125 ms), we observed a direct effect of age onto the response latency (model specifications: RMSEA = 0, CFI = 1, TLI = 1; saturated model; direct age effect; β = -0.141, SE = 0.066, z-value = -2.144, p = 0.032). Although age significantly impacts the PTA (β = 0.484, SE = 0.121, z-value = 4.003, p $<$ 0.001), there is no effect of PTA on the latency of this first peak in the spectrogram (β = -0.069, SE = 0.067, z-value = -1.030, p =0.303). Therefore, we do not observe a signifant indirect effect of age (indirect age effect; β = -0.033, SE = 0.033, z-value = -0.997, p = 0.319). This implies that the direct effect of age explains 81% of the observed total effect on the response latency of the first peak in the spectrogram.

Regarding the neural response latency of the second peak (150 to 250 ms) in the spectrogram, we observed

no significant impact of age in a direct or indirect way (model specifications: RMSEA = 0, CFI = 1, TLI = 1; saturated model). Age did not have an impact on the neural response latency (β = 0.031, SE = 0.060, z-value = 0.514, p = 0.607) nor did the PTA (β = 0.095, SE = 0.060, z-value = 1.591, p = 0.112), resulting in a non-significant indirect age effect (β = 0.046, SE = 0.031, z-value = 1.474, p = 0.140)

For the response latency of the first peak of the acoustic onsets (30 to 70 ms), we could differentiate the effect of age into its direct and indirect components (model specifications: RMSEA = 0, CFI = 1, TLI = 1; saturated model). The effect of age on the neural response latency did not reach significance (direct age effect: β = -0.038, SE = 0.027, z-value = -1.374, p =0.169). Similarly, no indirect effect of age, mediated through PTA was observed (indirect age effect: β = -0.014, SE = 0.013, z-value = -1.078, p =0.281). However, the total effect of age did reach significance (total age effect: β = -0.052, SE = 0.025, z-value = -2.095, p = 0.036). The latter indicates that there might be another measure, correlated with age, which explains observed effect of decreasing latency with increasing age. The direct age effect, however, explains 73% of the observed total effect on the response latency of the first peak in the acoustic onsets.

Regarding the effect of age on the response latency of the second peak of acoustic onsets, we did not observe any effect of age on the neural response latency model specifications: RMSEA = 0, CFI = 1, TLI = 1; saturated model). Age did not have an impact on the neural response latency (β = 0.004, SE = 0.051, z-value = 0.074, p = 0.941). PTA values increase with increasing age (β = 0.485, SE = 0.122, z-value = 3.986, p $<$ 0.001). PTA does not affect the neural response latency (β = 0.038, SE = 0.053, z-value = 0.719, p = 0.472).

6.2 Structural Equation Modelling - Linguistic

6.2.1 Neural Tracking

To investigate whether a decrease in hearing capacity or cognition explained the age effect observed across all electrodes and two later windows, we applied SEM. However, across all electrodes, we did not observe a mediation effect of age through SCWT, RST, or PTA (model specifications: RMSEA = 0, CFI = 1, TLI = 1.032). Age impacts linguistic NT directly (direct age effect: β = -0.358, SE = 0.171, z-value = -2.093, p = 0.036). Age also significantly impacts the SCWT (β = 0.493, SE = 0.121, z-value = 4.086, p $<$ 0.001), RST (β = -0.443, SE = 0.124, z-value = -3.565, p $<$ 0.001), and PTA (β = 0.478, SE = 0.122, z-value = 3.921, p $<$ 0.001). However, these predictors do not significantly impact linguistic NT (effect of SCWT on linguistic NT: β = -0.217, SE = 0.144, z-value =-1.507, p = 0.132; effect of RST on linguistic NT: β = -0.033, SE = 0.140, z-value = -0.236, p = 0.813; effect of PTA on linguistic NT: β = 0.207, SE = 0.143, z-value = 1.448, p = 0.148). This results in a total effect of age on linguistic NT is β = -0.352 (SE = 0.131, z-value = -2.693, p = 0.007).

When the linguistic NT values are averaged across centro-parietal electrodes and these two later time

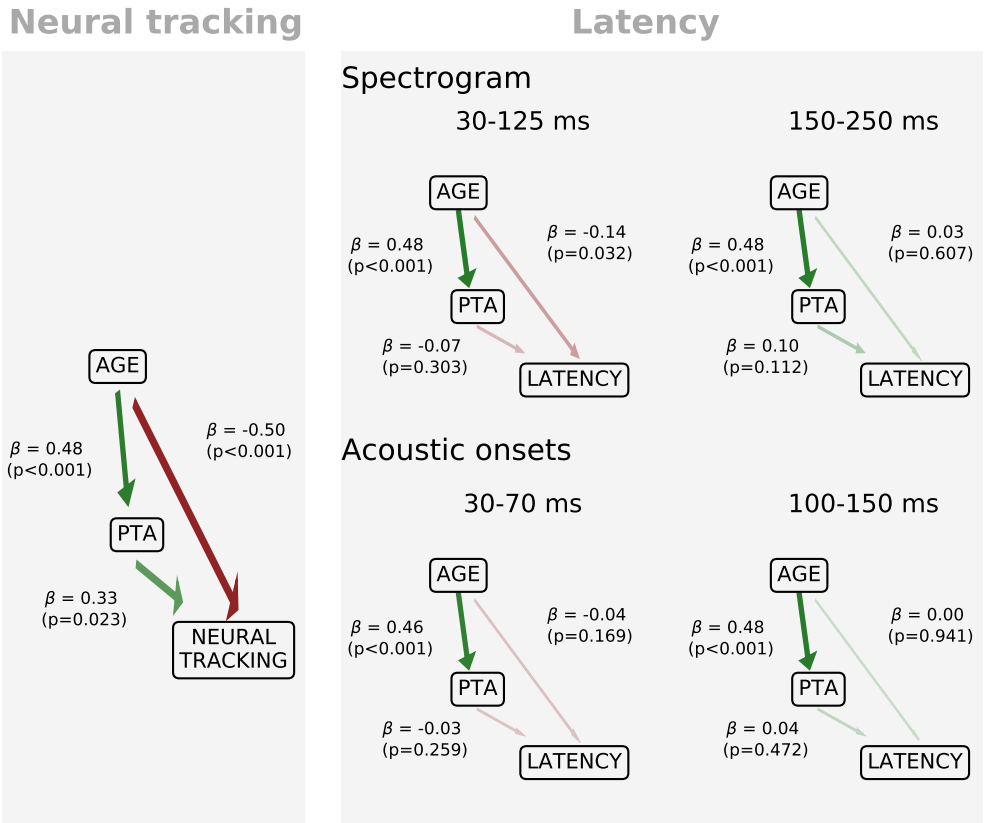

Figure S.1: Results obtained with SEM for the acoustic model for neural tracking (left panel; for centro-parietal and temporal electrodes as indicated on fig. 1) and latency of the neural responses to the spectrogram (top, right panel) and acoustic onsets (bottom, right panel). The observed effects are shown with their estimate (represented by β) and their p-value. Green arrows indicate a positive effect, while red values indicate a negative effect.

windows, the same conclusion is drawn: the effect of age is not mediated through SCWT, RST, or PTA (model specifications: RMSEA = 0, CFI = 1, TLI = 1.032). Here, the direct age effect of age is β = -0.522 (SE = 0.164, z-value = -3.176, p = 0.001), while the total effect of age, i.e. direct and indirect age effects, is β = -0.467 (SE = 0.123, z-value = -3.802, p < 0.001).

In the later time window (300 to 600 ms), we observed significant impact of age on linguistic NT when averaged across all electrodes, centro-parietal electrodes, and temporal electrodes, but not when averaged across the frontal electrode selection (using a linear modeling approach). To investigate these effects in more detail, we applied SEM. For simplicity purposes, we fitted three models: (1) a SEM based on the linguistic NT values across all electrodes, (2) a SEM across centro-parietal and temporal electrodes, and (3) a SEM across frontal electrodes. These results are visualized in the left panel of fig. S.2.

However, in the late time window (300-600ms), averaged across all electrodes, we observed a direct age effect and an indirect effect of age through the SCWT scores (model specifications: RMSEA = 0, CFI = 1, TLI = 1.029). We observed a direct effect of age on linguistic NT (β = -0.347, SE = 0.157, z-value = -2.216, p = 0.027). No indirect effect of PTA and RST were observed. However, we did observe a significant indirect effect of age through the SCWT score(β = -0.196, SE = 0.081, z-value = -2.422, p = 0.015): the SCWT score increases with increasing age (β = 0.493, SE = 0.121, z-value = 4.086, p < 0.001), and a higher SCWT score decreases linguistic NT (β = -0.397, SE = 0.132, z-value = -3.007, p = 0.003). The total indirect effect of age on linguistic NT is β = -0.075 (SE = 0.120, z-value = -0.626, p =0.531; including the effect of SCWT, RST, and PTA). Altogether, this suggests that the age explains 82% of the observed total effects of age.

For the centro-parietal or temporal electrode selection, the effect of age on linguistic NT was not mediated by PTA, RST or SCWT (model specifications: RMSEA = 0, CFI = 1, TLI = 1.029). We observed a direct age effect (β = -0.496, SE = 0.081, z-value = -2.422, p = 0.015). However, no indirect effect of age through PTA, RST or SCWT reached significance. The total indirect of effect of age is β = -0.018 (SE = 0.112, z-value = -0.161, p = 0.872). The direct age affect therefore explains 97% of the observed total age effect.

The indirect age effect mediated through SCWT scores was only observed when the prediction accuracies were averaged across all electrodes. Neither the temporal or centro-parietal electrode selection could explain this. Therefore, we also analysed the frontal electrode selection. Indeed, in this frontal electrode selection, although we did not observe a significant direct age effect, we observed a significant indirect effect of age mediated through the SCWT scores (model specifications: RMSEA = 0, CFI = 1, TLI = 1.038; indirect effect: β = -0.186, SE = 0.086, z-value = -2.158, p = 0.031). The direct age effect becomes non-significant, which align with the results of the above analysis (β = 0.113, SE = 0.176, z-value = 0.639, p = 0.523). With increasing age, the SCWT scores increase (β = 0.493, SE = 0.121 z-value = 4.086, p < 0.001), which leads to a decrease in frontal linguistic NT (β = -0.377, SE = 0.148, z-value = -2.541, p

= 0.011).

6.2.2 Neural Response Latency

As described in the method section, we investigated whether age impacted the latency of the neural responses for each linguistic speech representation. However, using a linear model, we only observed age effects on the latency of the later response to cohort entropy and the later response of word surprisal. Using a linear model, we did not find any effects of age on the early responses of linguistic speech representations or the peak latencies in the neural response to phoneme surprisal.

As expected, the SEM models converge with the results of the linear models: no significant direct nor total indirect effects of age were found for early responses of linguistic speech representations or the peak latencies in the neural response to phoneme surprisal. For simplicity purposes, these models are therefore not discussed in more detail.

For the latency of the later response to cohort entropy (250 to 350 ms), the SEM analysis indicated a significant direct effect of age on the neural response latency (model specifications: RMSEA = 0, CFI = 1, TLI = 1.033; direct effect: β = 0.121, SE = 0.053, z-value = 2.266, p = 0.023). This effect suggests that as age increases, the latency of the later neural response to cohort entropy increases. However, interestingly, also a significant indirect effect of age through PTA is observed on this latency (indirect age effect through PTA: β = -0.084, SE = 0.031, z-value = -2.740, p = 0.006). Merely, with increasing age the PTA values increase as well (β = 0.466, SE = 0.130, z-value = 3.594, p $<$ 0.001). Furthermore, as these PTA increases the latency of the later response to cohort entropy decreases (β = -0.181, SE = 0.043, z-value = -4.233 p $<$ 0.001). These results suggest that the effect of age and the effect of hearing loss affect the latency of the later response to cohort entropy in opposite directions.

Although we did not find any significant direct age effect or total indirect age effect, we would also like to bring the SEM analysis for the early activity to word surprisal (125 to 255 ms) to the reader's attention (model specifications: RMSEA = 0.09, CFI = 0.973, TLI = 0.912). In this analysis, we observed a marginal significant indirect effect of age through RST on the early neural response latency to word surprisal (indirect age effect through RST: β = -0.057, SE = 0.028, z-value = -1.988, p = 0.047). This effect indicates that as age increases, the RST decrease (β = -0.509, SE = 0.133, z-value = -3.827, p $<$ 0.001) which on its term increases the early neural response latency to word surprisal (β = 0.111, SE = 0.048, z-value = 2.326, p = 0.020). This indicates that there is a weak effect of age mediated through the RST on the latency of word surprisal.

In the linear model, we observed that age significantly increased the latency of the neural responses to the late negative peak to word surprisal (350 to 450 ms). However, interestingly, using the SEM analysis, we did not find a direct effect of age on this latency (model specifications: RMSEA = 0, CFI = 1, TLI = 1.036; direct age effect: β = 0.063, SE = 0.059, z-value = 1.059, p = 0.290). Additionally, neither did

we find a significant indirect effect of age mediated through PTA, RST or SCWT. This was unexpected as we did observe a significant effect of age on the latency of word surprisal. However, the SEM analysis indicated that higher SCWT is associated with increased latencies (β = 0.115, SE = 0.050, z-value = 2.278, p = 0.023). However, the associated indirect effect of SCWT through age did not reach significance (indirect age effect through SCWT: β = 0.051, SE = 0.027, z-value = 1.916, p = 0.055). This indicates that the effect of age observed in the linear model is likely more due to the correlation between age and SCWT rather than a robust effect of age. This SEM analysis indicated that the direct effect of age explains 51% of the observed total age effect which include direct and indirect effects (total age effect: β = 0.123, SE = 0.047, z-value = 2.610, p = 0.009).

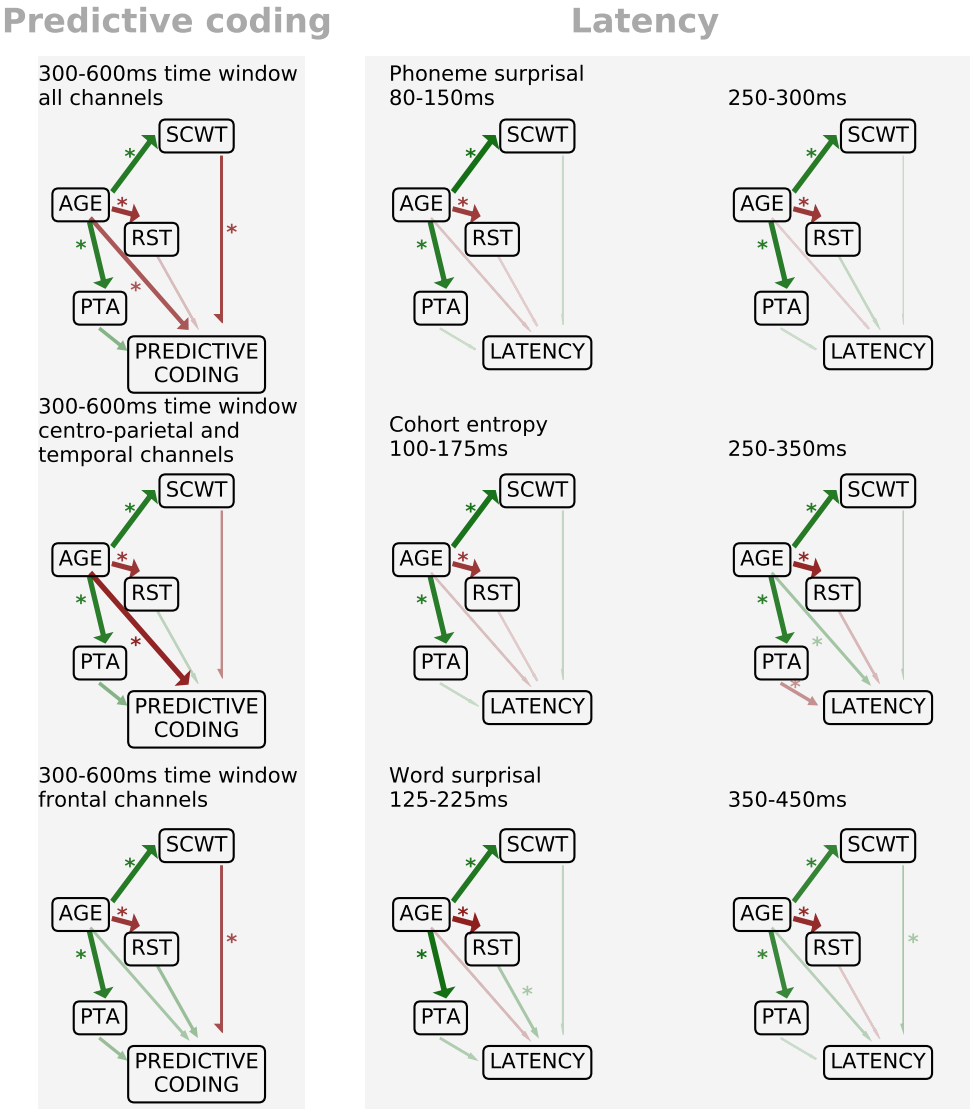

Figure S.2: Results obtained with SEM for the linguistic NT (left panel; each row indicates a different electrode selection; fig. 1) and latency of the neural responses to phoneme surprisal (top, right panel), cohort entropy (middle, right panel), and word surprisal (bottom, right panel). Significant estimates are annotated with an asterisk shown in the same color as the arrow. Green arrows indicate a positive effect, while red values indicate a negative effect. The darkness and width of the arrow indicate the magnitude of the effect.