Hearing loss is associated with delayed neural responses to continuous speech

Abbreviated title: Hearing loss delays neural responses to speech

Marlies Gillis[1], Lien Decruy[2], Jonas Vanthornhout[1], Tom Francart[1]

[1]: KU Leuven, Department of Neurosciences, ExpORL, 3000 Leuven, Belgium

[2]: Institute for Systems Research, University of Maryland, College Park, MD 20740, USA

Corresponding authors:

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

Tom Francart (tom.francart@kuleuven.be)

Abstract

We investigated the impact of hearing loss on the neural processing of speech. Using a forward modelling approach, we compared the neural responses to continuous speech of 14 adults with sensorineural hearing loss with those of age-matched normal-hearing peers.

Compared to their normal-hearing peers, hearing-impaired listeners had increased neural tracking and delayed neural responses to continuous speech in quiet. The latency also increased with the degree of hearing loss. As speech understanding decreased, neural tracking decreased in both population; however, a significantly different trend was observed for the latency of the neural responses. For normal-hearing listeners, the latency increased with increasing background noise level. However, for hearing-impaired listeners, this increase was not observed.

Our results support that the neural response latency indicates the efficiency of neural speech processing. Hearing-impaired listeners process speech in silence less efficiently then normal-hearing listeners. Our results suggest that this reduction in neural speech processing efficiency is a gradual effect which occurs as hearing deteriorates. Moreover, the efficiency of neural speech processing in hearing-impaired listeners is already at its lowest level when listening to speech in quiet, while normal-hearing listeners show a further decrease in efficiently when the noise level increases.

From our results, it is apparent that sound amplification does not solve hearing loss. Even when intelligibility is apparently perfect, hearing-impaired listeners process speech less efficiently.

Key words: neural tracking, hearing loss, speech, EEG

Introduction

It is widely known that hearing loss alters the brain (Eggermont, 2017; Peelle and Wingfield, 2016). To study the functional neural changes, several studies focussed on cortical auditory evoked potentials (CAEP) using electroencephalography (EEG). CAEPs reflect the cortical responses evoked by repetitions of simple sounds such as syllables, tone pips, or clicks. These responses represent the detection and/or discrimination of a sound. The CAEP-response is characterized by a first positive peak (P1) around 50 ms, a first negative peak (N1) around 100 ms and a later positive peak (P2) around 180 ms (Burkard et al., 2007). Harkrider et al. (2009) and Campbell and Sharma (2013) reported increased P2-latencies in hearing impaired listeners (HI listeners) compared to normal hearing listeners (NH listeners). Interestingly, Campbell and Sharma (2013) reported that P2-latency was also correlated with the person's speech perception ability in noise. Although changes in latency are often not reported, in most studies HI listeners showed increased amplitudes compared to NH listeners (Tremblay et al., 2003; Harkrider et al., 2006; Bertoli et al., 2011; Alain, 2014; Maamor and Billings, 2017) while Billings et al. (2015) and Koerner and Zhang (2018) did not observe differences between these two populations or others attributed these differences to decreased audibility of the stimulus (Oates et al., 2002; Van Dun et al., 2016; McClannahan et al., 2019). No consensus has been reached on the impact of hearing loss on the P1-N1-P2-complex. The use of continuous speech as the stimulus can be key to characterize the neural differences between these two populations as it requires more in-depth neural processing of the stimulus to understand the speech.

A limited number of studies has been conducted to study the effect of hearing loss on the neural responses to continuous speech. In these studies, the amount of neural tracking, i.e. to what extent speech is tracked by the brain, has been investigated in a two-talker scenario: an attended speaker and an ignored one (Petersen et al., 2017; Mirkovic et al., 2019; Presacco et al., 2019; Decruy et al., 2020; Fuglsang et al., 2020). In all these studies, both NH listeners and HI listeners, showed a higher neural tracking of the attended speech stream than that of the ignored speech stream. Petersen et al. (2017) reported that adults with a higher degree of hearing loss showed a higher neural tracking of the ignored speech and no change in the attended stream, suggesting that they experience more difficulties inhibiting irrelevant information. Although Mirkovic et al. (2019) and Presacco et al. (2019) did not report a neural difference between the two populations, Decruy et al. (2020) and Fuglsang et al. (2020) observed, in contrast to Petersen et al. (2017), an enhanced neural tracking in HI listeners for the attended-speech compared to their normal-hearing peers. This enhancement can indicate a compensation mechanism: HI listeners need to compensate for the degraded auditory input and therefore show increased cortical neural responses.

The difficulties of researching HI listeners are twofold. First, most HI listeners are older, and ageing also has an impact on brain responses (Tremblay et al., 2003; Harkrider et al., 2006; Burkard et al., 2007; Harkrider et al., 2009; Decruy et al., 2019; Presacco et al., 2016). Therefore, it is important to compare HI listeners to age-matched normal-hearing peers. Second, audibility of the stimulus must be taken into account: sound presented at the same intensity can be less audible for HI listeners than for NH listeners.

In the current study, we investigated whether there are differences in the neural responses to continuous speech between HI listeners and their age-matched normal-hearing peers. Our results showed delayed neural responses to continuous speech in HI listeners. We hypothesized that HI listeners recruit more brain regions to understand speech, which is reflected in enhanced neural tracking of speech as well as a delay of neural responses.

Materials and Methods

Participants

We used a dataset containing EEG of 14 HI listeners (8♀) with sensorineural hearing loss and 14 aged-matched normal-hearing peers (13♀) (between 21 and 82 years old). The data were collected in a previous study by Decruy et al. (2020). Inclusion criteria were: (1) having Dutch as a mother tongue, (2) having symmetrical hearing and (3) absence of medical conditions and learning disorders. A cognitive screening, the Montreal Cognitive Assessment (Nasreddine, 2004), was performed for all participants to ensure the absence of cognitive impairment. Hearing thresholds were determined using pure tone audiometry (125 to 8000 Hz). Normal hearing was defined for all participants where the hearing threshold did not exceed 30 dB HL for frequencies 125 to 4000 Hz (average of hearing thresholds within this frequency range in the stimulated ear is denoted as the pure-tone average (PTA)). The hearing thresholds and PTA are shown in Figure 1 (NH listeners: average PTA= 13.27 ± 5.60 dB HL, HI listeners: average PTA = 44.46 ± 10.54 dB HL).

Experimental Procedures

Behavioural Experiment: Flemish Matrix sentence test

The Matrix sentence test was performed to determine the participant's Speech Reception Threshold (SRT) in speech weighted noise (SWN). These Matrix sentences have a standard grammatical structure, consisting of a name, a verb, a numeral, a colour and an object (Luts et al., 2014). The SRT represents the signal-to-noise ratio (SNR) at which 50% of the presented words are recalled correctly.

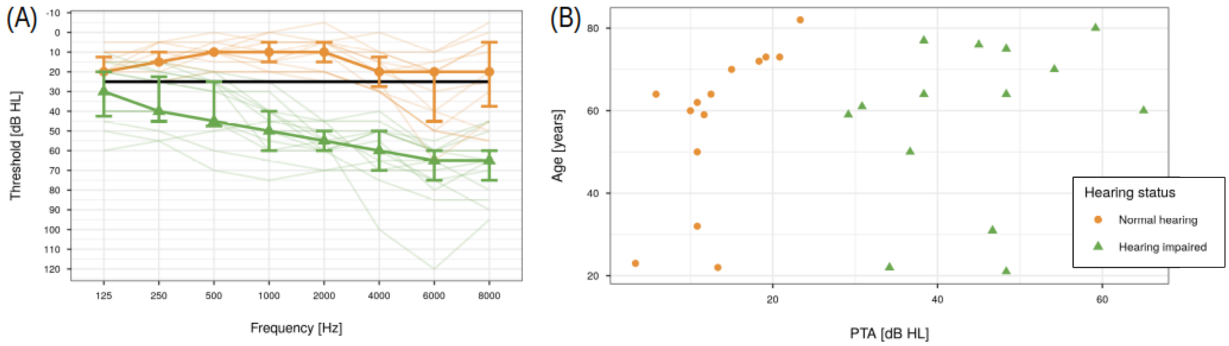

Figure 1: The hearing thresholds for the stimulated ear (panel A) and PTA as a function of age (panel B) for NH listeners (orange) and HI listeners (green).

EEG Experiment

Data acquisition A BioSemi ActiveTwo system (Amsterdam, Netherlands) was used to measure EEG signals during stimuli presentation. This system uses 64 Ag/AgCl electrodes placed according to the 10-20 system (Oostenveld and Praamstra, 2001). The EEG signals were measured with a sampling frequency of 8192 Hz. All recordings were carried out in a soundproof booth with Faraday cage at ExpORL (Dept. Neurosciences, KU Leuven).

Stimuli presentation The speech stimuli were presented monaurally through ER-3A insert phones (Etymotic Research Inc, IL, USA) using the software platform APEX (Dept. Neurosciences, KU Leuven) (Francart et al., 2008). The stimuli were presented to the right ear unless the participant preferred the left ear (n = 3; 1 NH; 2 HI). All stimuli were set to the same root mean square level and were calibrated.

For all NH listeners, the speech stimuli' intensity was fixed at 55 dB SPL (A-weighted). To ensure audible stimuli for HI listeners, the stimuli were linearly amplified based on the participant's hearing thresholds according to the National Acoustics Laboratory-Revised Profound (NAL) algorithm (Byrne et al., 2001). To ensure a comfortable level, the overall level was adjusted on a subject-specific basis in addition to the linear amplification so that the stimulus was minimally effortful and comfortable to listen to. The individual presentation levels are reported by Decruy et al. (2020).

During the EEG recording, 2 Dutch stories were presented: (1) "Milan", a 12-minute long story narrated by Stijn Vranken (♂) presented in quiet and (2) "De Wilde Zwanen" narrated by Katrien Devos (♀) presented in 5 different levels of background speech-weighted noise (each lasted around 2 minutes). The duration of silences was limited to 200 ms.

The levels of background noise for the second story depended on the participant's speech-in-noise performance. Using an adapted version of the self-assessed Békesy procedure (Decruy et al., 2019), the SRT of the Matrix sentences was adjusted to obtain a SRT of a story (Decruy et al., 2018, 2019). The noise conditions were calculated on the participant's story adjusted SRT, namely: SRT - 3 dB, SRT, SRT + 3 dB, SRT + 6 dB and a condition without noise, which approximate speech understanding levels of 20%, 50%, 80%, 95% and 100%. In addition to obtaining an estimate of the speech understanding in each noise condition, a subjective rating of the participant's speech understanding was obtained after each condition.

Signal Processing

Processing of the EEG signals

The EEG recording with a sampling frequency of 8192 Hz was downsampled to 256 Hz to decrease processing time. To remove artefacts of eye blinks, we applied multi-channel Wiener filtering to the EEG data to remove artefacts of eye blinks (Somers et al., 2018). Then we referenced the EEG data to the common-average and filtered the data between 0.5 and 25 Hz using a zero-phase Chebyshev filter (Type II with an attenuation of 80 dB at 10% outside the passband). Additional downsampling to 128 Hz was performed.

Extraction of the speech features

In this study, we used 2 speech features: spectrogram and acoustical onsets. Both variables are continuous speech features which represent the acoustical properties of the speech stimulus.

To create the spectrogram representation, the speech stimulus (without amplification) was low-pass filtered below 4000 Hz (zero-phase low-pass FIR filter with a hamming window of 159 samples) because the ER-3A insert phones also low-pass filter at this frequency. A spectrogram representation was obtained using the Gammatone Filterbank Toolkit 1.0 (Heeris, 2014) (centre frequencies between 70 and 4000 Hz with 256 filter channels and an integration window of 0.01 second). This toolkit calculates a spectrogram representation based on a series of gammatone filters inspired by the structure of the human auditory system (Slaney, 1998). The resulting 256 filter outputs were averaged into 8 frequency bands (each containing 32 outputs). Additionally, each frequency band was downsampled to the same sampling frequency as the processed EEG, namely 128 Hz. The NAL filtering introduced a delay of 5.334 ms, which was compensated for. The acoustical onsets representation was calculated as a half-wave rectification of the spectrogram's derivative.

Prediction accuracies, temporal response function & peak picking method

In this study, we focussed on a linear forward modelling approach which predicts the EEG based on a linear combination of speech features of the presented speech. This forward modelling approach results in 2 outcomes: (a) a temporal response function (TRF) and (b) a prediction accuracy. (a) A TRF is a linear approximation of the brain's impulse response. It is a signal over time which describes how the brain responds to the speech features. (b) TRFs can be used to predict the EEG by convolving it with the speech features. The predicted EEG is then correlated with the actual EEG to obtain a prediction accuracy. Prediction accuracy is considered a measure of neural tracking: the higher the prediction accuracy, the better the brain tracks the stimulus.

(a) To estimate TRFs, we used the Eelbrain toolbox (Brodbeck, 2020). The toolbox estimates TRFs using the boosting algorithm by David et al. (2007) (using a fixed step size of 0.005; stopping criteria based on ℓ_2-norm; kernel basis of 50 ms). We used 4-fold cross-validation (4 equally long folds; 3 folds used for training, 1 for validation) and an integration window between 0 and 700 ms. The estimated TRFs, averaged across folds and frequency bands, were used to determine the peak latencies.

(b) To calculate the prediction accuracy, the TRF is applied to left-out EEG to allow a fair comparison between models with a different number of speech features. We used the boosting algorithm with a testing fold. This implies a 4-fold cross-validation with 2 folds for training, 1 fold for validation and 1 fold for testing, which is left-out during training and validation. Each estimated TRF was used to predict the EEG of the left-out testing fold. The predicted EEG of all left-out segments are correlated, using Pearson correlation, with the actual EEG to obtain a prediction accuracy per EEG-electrode. The prediction accuracies were averaged across EEG-electrodes and denoted as neural tracking. Similarly, as Decruy et al. (2020), we calculated the neural tracking of the second story, presented in different level of background noise, using the TRFs estimated on the story in quiet.

From the TRF, we aimed to identify the amplitude and latency of 3 peaks: P1, N1 and P2. As the EEG data contains 64 different channels, 64 different TRFs were estimated, which made peak picking more complex. Therefore we applied principal component analysis (PCA), a dimensionality reduction method. The PCA-method results in (a) signals in component space and (b) corresponding spatial filters which describe the linear combinations of EEG channels to obtain these components. In our analysis, the first component was used. Adding more components up to 4 did not change the findings of this study. In addition to the time course of the component, we also investigated the corresponding spatial filter. As the sign of this spatial filter is arbitrary, we forced the average of occipital and parietal channels (P9, P7, PO7, O1, Oz, O2, PO8, P8, Iz, P10) to be negative by multiplying the spatial filter with -1 when needed. The PCA-method was applied to the data per story for each participant.

To identify the different peaks, we performed a z-score normalization of the TRF in component space and determined the maximal or minimal amplitude for positive and negative peaks in different time regions (P1: 30 to 110 ms, N1: 30 to 210 ms, P2: 110 to 270 ms), respectively. The overlap of these time regions is not an issue as we identified either the maximal or minimal amplitude to determine the peak latency of a positive or negative peak, respectively. To only identify prominent peaks, a peak was discarded from the analysis if the amplitude of the normalized TRF was smaller than the threshold of 1.

Statistical analysis

We used the R software package (version 3.6.3) (R Core Team, 2020). We used the Buildmer toolbox, which allows identifying the best linear mixed model (LMM) given a series of predictors and all their possible interactions based on the likelihood-ratio test (Voeten, 2020). We used the following predictors: (a) hearing status (NH or HI) or the PTA depending on whether we were interested in the group effect or the effect of the degree of hearing loss, (b) age and (c) peak type (P1, N1, P2). To observe an effect of model choice on prediction accuracy, we also included the predictor (d) model type (Spectrogram, Acoustic onsets, Acoustic onsets + Spectrogram) in the statistical analysis. The analysis over different noise conditions also included the predictor (e) speech understanding. For the latter analysis, all continuous predictors were z-scored to minimize effects due to differences in scale. A matching factor indicated the participants belonging to the same age-matched pair. We included a nested random effect: participant nested inside match, as each match contained a pair of participants, and each participant had multiple dependent observations. The models' assumptions were checked with a visual inspection of the residual plots to assure homoscedasticity and normality. The models' outcomes were reported with the unstandardized regression coefficient (β) with standard error (SE), t-value and p-value per fixed effect. If significant interaction effects were found or if we aimed to identify differences between different levels of a factor, additional Holm-adjusted posthoc tests were performed. A significance level of $\alpha = 0.05$ was set for all estimated models.

To compare differences in spatial filters or topographies of the peaks between the 2 groups, we applied a method proposed by McCarthy and Wood (1985) to determine whether the topography, when amplitude differences are discarded by normalization, depends on either the considered population or speech feature. A difference in topography or spatial filter is observed when there is a significant interaction between the sensor and condition, e.g., the considered population or speech feature. To test whether the N1 topography depended on the speech feature, we only included participants who showed a prominent peak for both speech features.

To determine whether the topography or spatial filter depended on the speech feature, the participant was included as a random effect. The p-value of the interaction term is reported. A significance level of $\alpha = 0.05$ was used.

Results

Neural differences when listening to speech in quiet

Hearing-impaired listeners show higher neural tracking

We identified the speech feature(s) that resulted in the highest neural tracking: acoustic onsets, spectrogram or a combination of both speech features. As shown in Figure 2 and verified by the statistical analysis, the highest neural tracking was obtained with a combination of both speech features (analysis using LMM: Table 1). Additionally, HI listeners showed higher neural tracking compared to the group of NH listeners (on average 0.017 higher, $p = 0.045$). Age did not have a significant effect on the neural tracking of speech. Post-hoc tests confirmed that the highest neural tracking was obtained with a combination of both speech features which was higher compared to the model using the acoustic onsets ($p = 0.019$) and higher compared to the model using the spectrogram variable ($p < .001$).

Table 1: Liner mixed model: the effect of hearing status and model type on neural tracking. Estimates of the regression coefficients (β), standard errors (SE), degrees of freedom (df), t-Ratios and p-values are reported per fixed-effect term. Participant nested in match was included as a random effect.
Formula: neural tracking $\sim$ 1 + hearing status + model type + (1 | match/participant)

Fixed-effect term	β	SE	df	t-Ratio	p-value
Intercept (for NH / Spectrogram)	0.029	0.006	26.396	4.918	p < .001
Hearing status: HL	0.017	0.008	26	2.106	p = 0.045
Model type = Acoustic onsets	0.003	0.001	54	3.937	p < .001
Model type = Acoustic onsets + spectrogram	0.006	0.001	54	6.357	p < .001

Delayed peak latencies for hearing-impaired listeners

In Figure 4.A, all TRFs in component space are shown for both populations and speech features. The TRFs of HI listeners show delayed neural responses to speech compared to those of normal-hearing adults. Additionally, the average TRF for both speech features show only 2 prominent peaks: P1-peak of acoustic onsets ($P1_{AO}$), N1-peak of acoustic onsets ($N1_{AO}$), N1-peak of spectrogram ($N1_S$), P2-peak of spectrogram ($P2_S$). The best LMM predicting latency included a main effect of factor peak, hearing status and age (Table 2). Adults with hearing loss showed later peak latencies (an increase of 22 ms, $p = 0.001$). The effect of age depended on the considered peak. No significant interaction between age and hearing status was

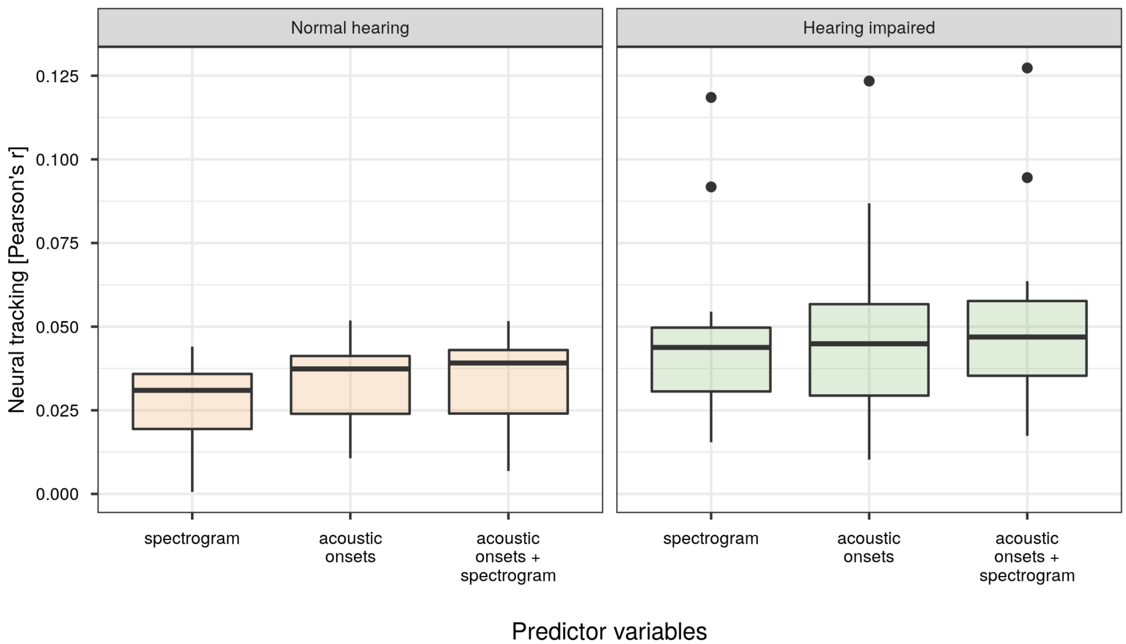

Figure 2: Neural tracking (Pearson's r) as a function of different combinations of speech features ('spectrogram', 'acoustic onsets' and 'acoustic onsets + spectrogram', respectively) for both NH listeners (left; orange) and HI listeners (right; green).

observed. Post-hoc testing showed a significant increase in latency with increasing age for the $P2_S$-latency ($p = 0.029$) while this was not observed for the other peak latencies ($P1_{AO}$: $p = 0.398$, $N1_{AO}$: $p = 0.398$, $N1_S$: $p = 0.398$).

We did not observe a significant difference between the spatial filters of HI listeners and NH listeners (for spectrogram: $p = 0.514$, for acoustic onsets: $p = 0.974$; Figure 3.A). HI listeners showed a significantly different topography for $N1_S$ and $P2_S$ compared to NH listeners ($N1_S$: $p < 0.001$; $P2_S$: $p < 0.001$, respectively) while this was not the case for $P1_{AO}$ ($p = 0.578$) and $N1_{AO}$ ($p = 0.145$) (Figure 3.B). The method by McCarthy and Wood (1985) does not allow to pinpoint the differences in topography between the two population. However, from Figure 3.B, we observed that HI listeners showed more left lateralized activity than NH listeners.

Table 2: Results of the linear mixed model in order to assess peak type, hearing status and age on the peak latency of $P1_{AO}$, $N1_{AO}$, $N1_S$, $P2_S$. Estimates of the regression coefficients (β), standard errors (SE), degrees of freedom (df), t-Ratios and p-values are reported per fixed-effect term. Participant nested in match was included as a random effect.
Formula: latency $\sim$ 1 + peak + hearing status + age + peak:age + (1 | match/participant)

Fixed-effect term	β	SE	df	t-Ratio	p-value
Intercept (for NH / for P1 - acoustic onsets)	65.427	12.964	57.639	5.047	p < .001
peak = N1 - acoustic onsets	77.141	15.661	61.848	4.926	p < .001
peak = N1 - spectrogram	32.329	15.917	62.021	2.031	p = 0.047
peak = P2 - spectrogram	94.394	16.084	62.433	5.869	p < .001
Hearing status: HL	22.428	5.928	23.694	3.783	p = 0.001
Age	-0.322	0.211	61.937	-1.525	p = 0.132
peak = N1 - acoustic onsets:age	-0.02	0.253	61.221	-0.078	p = 0.938
peak = N1 - spectrogram:age	0.118	0.255	61.14	0.462	p = 0.646
peak = P2 - spectrogram:age	0.988	0.263	62.12	3.76	p < .001

Longer latencies are associated with higher degrees of hearing loss

As significant differences in peak latencies were observed between the two populations, we hypothesized that a higher degree of hearing loss is associated with increased latency of the peaks. The best linear mixed model to predict latency contained main effects of factor peak, degree of hearing loss, age and the interaction between age and peak type. Post-hoc testing did not show significant effects of age on the peak latencies, so this effect was not visualized in Figure 4. With increasing degree of hearing loss, the latency of neural responses increased ($\beta = 0.712$, $p < .001$; Table 3; Figure 4.B).

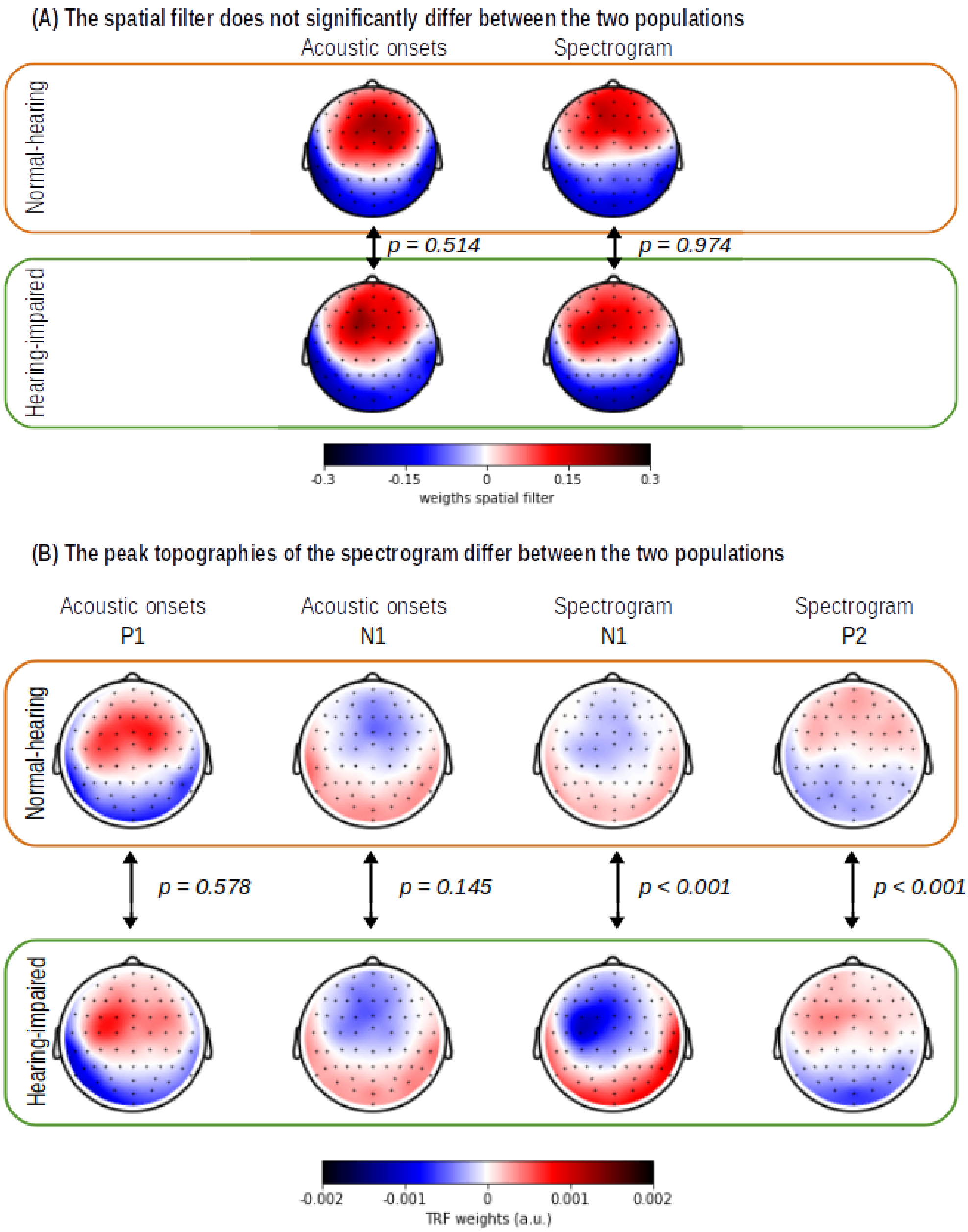

Figure 3: Visualization of the PCA spatial filters (panel A) and the topographies of the peaks (panel B) in the TRFs in sensor space for both speech features, spectrogram and acoustic onsets, and for the two populations: NH listeners and HI listeners.

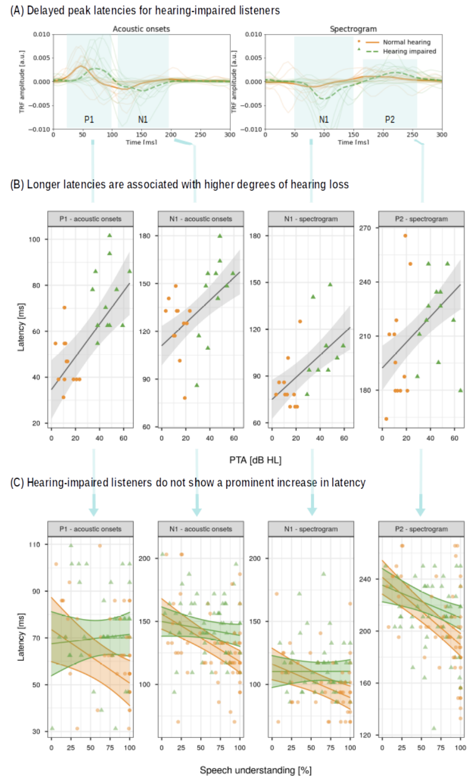

Figure 4: An overview of the neural responses of HI listeners (HI; striped line; green, triangle) and NH listeners (NH; orange, dot). Panel A: TRF in component space when listening to a story in quiet for both speech features evaluated for both populations. The thick line represents the average TRF over participants. The lighter lines represent the subject-specific TRFs. Panel B: The peak latency in function of the degree of hearing loss (PTA) derived from the neural responses when listening to a story in quiet. Panel C: The peak latencies in function of speech understanding derived from the neural responses when listening to a story presented in multiple levels of background noise. The effect of degree of hearing loss was made discrete at 2 levels, the average hearing thresholds of all NH listeners (level NH listeners: 13 dB HL) and subjects with hearing loss (level HI listeners: 44 dB HL) and is represented by the regression lines with confidence intervals (shaded area).

Table 3: Results of the linear mixed model in order to assess the effects of degree of hearing loss and age on the peak latency of $P1_{AO}$, $N1_{AO}$, $N1_S$, $P2_S$. Estimates of the regression coefficients (β), standard errors (SE), degrees of freedom (df), t-Ratios and p-values are reported per fixed-effect term. Participant nested in the matching factor was included as a random nested effect.
Formula: latency $\sim$ 1 + peak + degree of hearing loss + age + peak:age + (1 | match/participant)

Fixed-effect term	β	SE	df	t-Ratio	p-value
Intercept (for P1 - acoustic onsets)	62.136	12.619	60.398	4.924	$p < .001$
peak = N1 - acoustic onsets	76.879	15.67	61.756	4.906	$p < .001$
peak = N1 - spectrogram	33.13	15.924	61.914	2.08	$p = 0.042$
peak = P2 - spectrogram	94.424	16.087	62.406	5.87	$p < .001$
Degree of hearing loss (PTA)	0.712	0.161	23.625	4.412	$p < .001$
Age	-0.43	0.206	64.874	-2.085	$p = 0.041$
peak = N1 - acoustic onsets:age	-0.008	0.254	61.035	-0.032	$p = 0.974$
peak = N1 - spectrogram:age	0.113	0.255	60.984	0.441	$p = 0.661$
peak = P2 - spectrogram:age	0.988	0.263	62.052	3.762	$p < .001$

Neural differences when speech understanding decreases

Increased neural tracking with increased speech understanding

The effect of increased neural tracking for HI listeners was robust over different levels of background noise ($\beta = 1.65\text{e-}02$, $p = 0.001$; Table 4; Figure 5). Additionally, higher neural tracking was observed with increasing age ($\beta = 3.52\text{e-}04$, $p = 0.006$; Table 4; Figure 5) and with increasing speech understanding ($\beta = 1.54\text{e-}04$, $p < .001$; Table 4; Figure 5). No significant interaction effect was observed between hearing status and speech understanding.

Table 4: Linear mixed model: the effect of hearing status and speech understanding on neural tracking. Estimates of the regression coefficients (β), standard errors (SE), degrees of freedom (df), t-Ratios and p-values are reported per fixed-effect term. Participant nested in match was included as a random effect. Formula: neural tracking $\sim$ 1 + hearing status + age + speech understanding + (1 | match/participant)

Fixed-effect term	β	SE	df	t-Ratio	p-value
Intercept (for NH)	-0.011	0.008	29.734	-1.464	p = 0.154
Hearing status: HL	1.65e-02	4.44e-03	25.175	3.721	p = 0.001
Age	3.52e-04	1.18e-04	24.988	2.981	p = 0.006
Speech understanding	1.54e-04	2.96e-05	115.604	5.185	p < .001

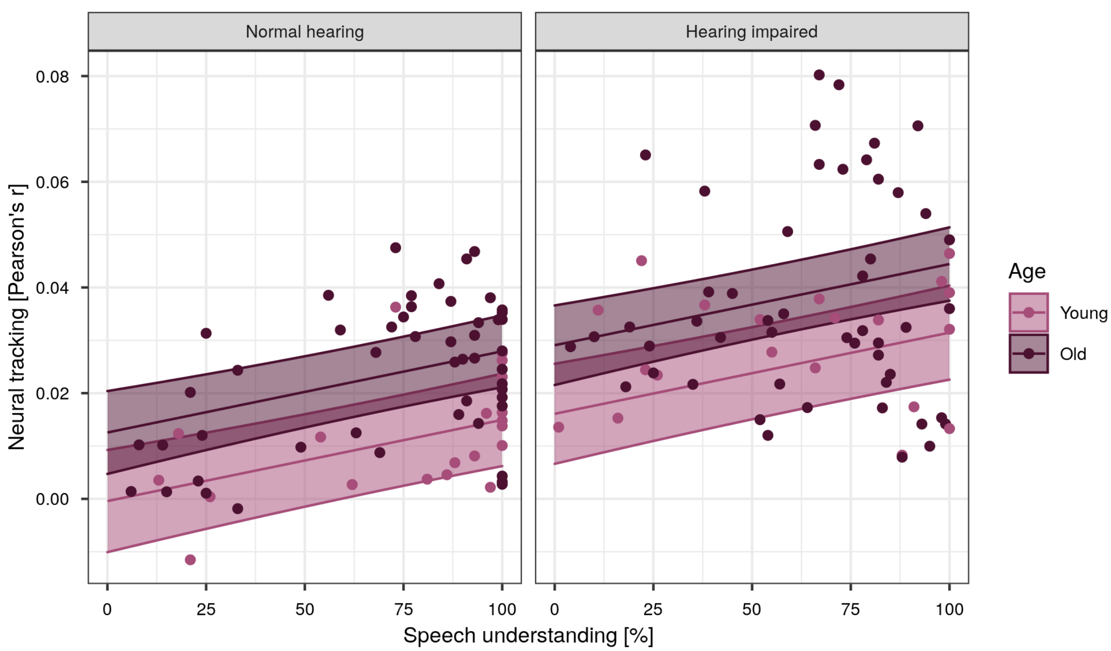

Figure 5: Neural tracking (Pearson's r) as a function of speech understanding. The effect of age is visualized by 2 regression lines with confidence intervals (shaded area) for a sample person aged 31 (young; pink) and 68 (old; purple).

Hearing-impaired listeners do not show a prominent increase in latency

We analysed the effects of speech understanding, age and degree of hearing loss on the peak latencies for the second story presented in different levels of background noise. The trends of age and speech understanding on peak latency depended on the considered peak. However, posthoc tests did not show a significant effect of age on the latencies of the peaks. Therefore this effect is not visualized in Figure 4.C. Interestingly, we found a significant interaction effect between speech understanding and degree of hearing loss ($p < .001$; Table 5; Figure 4.C). Post-hoc testing showed that NH listeners showed a significant increase in latency when speech understanding decreased ($p < .001$) while no significant increase was observed for HI listeners ($p = 0.34$). This trend significantly differed between the two populations ($p < .001$).

Table 5: Results of the linear mixed model in order to assess the effects of degree of hearing loss, speech understanding and age on the peak latency of $P1_{AO}$, $N1_{AO}$, $N1_S$, $P2_S$. Estimates of the regression coefficients (β), standard errors (SE), degrees of freedom (df), t-Ratios and p-values are reported per fixed-effect term. Participant nested in the matching factor was included as a random nested effect.
Formula: latency ~ 1 + peak + SI + degree of hearing loss + speech understanding:degree of hearing loss + peak:speech understanding + age + degree of hearing loss:age + peak:age + speech understanding:age + (1 | match/participant)

Fixed-effect term	β	SE	df	t-Ratio	p-value
Intercept (for P1 - acoustic onsets)	65.523	3.149	69.524	20.808	p < .001
peak = N1 - acoustic onsets	71.012	3.015	384.446	23.557	p < .001
peak = N1 - spectrogram	39.875	3.203	385.734	12.449	p < .001
peak = P2 - spectrogram	144.887	3.097	385.778	46.783	p < .001
Speech understanding	-2.46	2.336	388.475	-1.053	p = 0.293
Degree of hearing loss (PTA)	5.489	2.324	23.74	2.362	p = 0.027
Age	-6.421	2.882	59.27	-2.228	p = 0.03
Speech understanding:Degree of hearing loss (PTA)	4.441	1.071	396.822	4.148	p < .001
peak = N1 - acoustic onsets:speech understanding	-4.487	2.968	381.779	-1.512	p = 0.131
peak = N1 - spectrogram:speech understanding	-1.148	3.163	382.021	-0.363	p = 0.717
peak = P2 - spectrogram:speech understanding	-8.703	3.066	383.334	-2.839	p = 0.005
Degree of hearing loss (PTA):Age	4.797	2.232	24.283	2.149	p = 0.042
peak = N1 - acoustic onsets:age	4.906	2.832	379.405	1.732	p = 0.084
peak = N1 - spectrogram:age	0.94	3.196	384.574	0.294	p = 0.769
peak = P2 - spectrogram:age	13.103	2.897	384.584	4.523	p < .001
Speech understanding:Age	-2.121	1.082	390.971	-1.96	p = 0.051

The effect of the degree of hearing loss on the peak amplitude was not consistent for all peaks.

Discussion

We compared the neural responses to continuous speech of adults with a sensorineural hearing loss with those of age-matched normal-hearing peers. We found that HI listeners show higher neural tracking and increased peak latencies in their neural responses. Across noise conditions, NH listeners showed increased latencies as speech understanding decreased. However, for adults with hearing loss, this increase in latency was not observed.

Higher neural tracking of speech in hearing-impaired listeners

By evaluating neural tracking, we concluded that (1) higher neural tracking is observed for a combination of the spectrogram and acoustic onsets compared to the speech features individually and (2) HI listeners show enhanced neural tracking compared to normal-hearing peers.

The combination of speech features results in higher neural tracking, which implies that both speech features encode different information. Following Hamilton et al. (2018) and Brodbeck et al. (2020), both speech features allow a differentiation between sustained activity, represented by the spectrogram, and transient activity, represented by acoustic onsets. When the TRF is estimated using only the spectrogram, it shows a positive peak around 50 ms (P1), negative peak around 100 ms (N1) and a positive peak between 100 ms and 200 ms P2 (Di Liberto et al., 2015; Lesenfants et al., 2019b). In our data, the TRF shows only 2 prominent peaks for each speech feature. We hypothesize that the P1-peak is mainly dominated by transient activity while the P2-peak is dominated by sustained activity. Although the difference in latency, the N1-peak is observed in the TRF of both speech features and resulted in a similar topography. This suggests that the N1-response is evoked by spatially similar neural sources which respond differently to speech.

Using a linear modelling approach, Decruy et al. (2020) and Fuglsang et al. (2020) also reported that HI listeners have higher neural tracking than NH listeners of the attended speaker. Nevertheless, Presacco et al. (2019) did not find a difference in neural tracking between the two populations. However, in their study, the populations were not closely age-matched, while ageing is known to increase neural tracking (Presacco et al., 2016; Decruy et al., 2019).

Like previous literature, we also observed that neural tracking decreases with decreasing speech understanding (Vanthornhout et al., 2018; Lesenfants et al., 2019a; Decruy et al., 2020). Even when the speech is presented with background noise, HI listeners showed enhanced neural tracking of speech. This suggests evidence for a compensation mechanism: higher neural tracking indicates more neural activity to compensate for the degraded auditory input (Eggermont, 2017; Fuglsang et al., 2020). Although the enhanced neural tracking of speech in HI listeners, the effect of speech understanding on the neural tracking was similar for both populations.

Hearing-impaired listeners process speech less efficiently

HI listeners showed significantly increased latencies compared to their age-matched normal-hearing peers when they listened to a story presented in quiet (Figure 4.A). Additionally, the delay in neural responses increased with a higher degree of hearing loss (Figure 4.B).

We are not aware of any reports in which the effect of hearing loss on the latency of neural responses is studied using continuous speech presented in quiet and with background noise. However, investigating the CAEP-response, Campbell and Sharma (2013) and Bidelman et al. (2019b) have reported an increased P2 latency with worse speech perception in noise but not with the degree of hearing loss. However, in both studies, the same intensity was presented to both populations, therefore McClannahan et al. (2019) remarked that differences in the audibility of the stimulus might explain the differences in neural response latency. Recently, Verschueren et al. (2020) observed that reduced audibility increases the latency of the neural responses to continuous speech. However, at a comfortable loudness (at intensities of 60 dB or higher in a NH population), the latency reaches a plateau. As our stimulus is amplified based on the participants' hearing thresholds and is presented at a subject-specific intensity to assure comfortable listening for HI listeners, we minimized the effects of differences in audibility of the stimulus.

In several studies, it has been shown that increasing task demand due to lower stimulus intensity or increasing background noise is associated with an increase in the latency of neural responses in continuous speech (Mirkovic et al., 2019; Verschueren et al., 2020) or in CAEP-responses (Billings et al., 2015; Van Dun et al., 2016; Maamor and Billings, 2017; McClannahan et al., 2019). Similarly, for NH listeners, we found that as speech understanding decreases, the delay of neural responses increases. However, this increase in latency is absent for adults with a higher degree of hearing loss (Figure 4.C). This could explain why Mirkovic et al. (2019) did not find a difference between the two populations since they presented only two noise conditions. As the noise level increases, the difference in latency between the two populations becomes smaller, which reduces the likelihood of a statistical difference between the two populations.

In addition to an increase in latency of the P2 peak, Campbell and Sharma (2013) and Bidelman et al. (2019b) also reported increased frontal activation in HI listeners in the neural responses to simple sounds. This converges with our results using continuous speech: HI listeners show more left-lateralized frontal activity for the $N1_S$ and $P2_S$ topographies (Figure 3). This suggest that HI listeners recruit different underlying neural sources to understand speech.

Similar to the findings of Campbell and Sharma (2013), Bidelman et al. (2019a) and Bidelman et al. (2019b), our results indicate that the latency of the neural responses reflects the efficiency of neural processing: if more or different brain regions are involved to process the speech, suggested by the difference in topography. Hence it causes longer communication pathways in the brain (Bidelman et al., 2019b) which increases the processing time, show by the difference in the latency and reduces the efficiency of neural speech processing. Therefore, we suggest latency as a marker for the efficiency of neural processing of continuous, natural speech.

The latencies of the neural responses when listening to speech in quiet are significantly increased for HI listeners. This means that hearing loss causes less efficient processing of speech although the sound is amplified based on the listener's hearing thresholds. Furthermore, the efficiency of neural processing decreases as the severity of the hearing loss increases. Therefore, we hypothesize that the recruitment of additional brain regions and the corresponding decrease in efficiency of neural speech processing, is a gradual effect which occurs as the hearing deteriorates.

Bidelman et al. (2019b) did not report an effect of noise on the P2-latency. However, investigating the functional brain connectivity in the same data, Bidelman et al. (2019a) reported that as noise was added to the stimulus, NH listeners showed more long-range neural signalling whereas this was not seen for HI listeners (Bidelman et al., 2019a). The latter finding is supported by our data: in NH listeners the neural response latency increases as the speech understanding decreases due to increasing level of background noise. This suggests less efficient neural processing to understand speech in noise: more processing time is required to attend the speech stream and ignoring the noise. However, for HI listeners, this is not the case: when background noise increases, processing efficiency does not decrease. It can be that HI listeners cannot recruit additional brain regions in the speech network as they already recruit a maximum number of brain regions in the speech network to understand speech in quiet.

Finally, we would like to highlight the difference in the trend of neural tracking and neural response latency. As speech understanding decreases, neural tracking decreases for both NH and HI listeners while the neural response latency remains constant (HI) or increases (NH). This difference in trend suggests that both measures represent different underlying neural processes for speech comprehension.

Conclusion

In this study, we compared the neural responses to continuous speech of adults with a sensorineural hearing loss with those of age-matched normal-hearing peers. HI listeners showed increased peak latencies of their neural responses. Interestingly, the latency increases as the degree of hearing loss increases. Across noise conditions, latency generally increases as the listening conditions become more difficult. However, for HI listeners, this increase in latency is not observed. We here suggest latency as a marker for the efficiency of neural processing to understand continuous, natural speech.

References

Alain, C. (2014). Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. *Frontiers in systems neuroscience*, 8:8.

Bertoli, S., Probst, R., and Bodmer, D. (2011). Late auditory evoked potentials in elderly long-term hearing-aid users with unilateral or bilateral fittings. *Hearing research*, 280(1-2):58–69.

Bidelman, G. M., Mahmud, M. S., Yeasin, M., Shen, D., Arnott, S. R., and Alain, C. (2019a). Age-related hearing loss increases full-brain connectivity while reversing directed signaling within the dorsal–ventral pathway for speech. *Brain Structure and Function*, 224(8):2661–2676.

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

Billings, C. J., Penman, T. M., McMillan, G. P., and Ellis, E. (2015). Electrophysiology and perception of speech in noise in older listeners: effects of hearing impairment & age. *Ear and hearing*, 36(6):710.

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

Brodbeck, C., Jiao, A., Hong, L. E., and Simon, J. Z. (2020). Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *bioRxiv*, page 866749.

Burkard, R. F., Eggermont, J. J., and Don, M. (2007). *Auditory evoked potentials: basic principles and clinical application*. Lippincott Williams & Wilkins.

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (2001). Nal-nl1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *Journal of the American academy of audiology*, 12(1).

Campbell, J. and Sharma, A. (2013). Compensatory changes in cortical resource allocation in adults with hearing loss. *Frontiers in systems neuroscience*, 7:71.

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., Das, N., Verschueren, E., and Francart, T. (2018). The self-assessed Békesy procedure: validation of a method to measure intelligibility of connected discourse. *Trends in hearing*, 22:2331216518802702.

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*, page 107961.

Di Liberto, G. M., O'Sullivan, J. A., and Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, 25(19):2457–2465.

Eggermont, J. J. (2017). Acquired hearing loss and brain plasticity. *Hearing Research*, 343:176–190.

Francart, T., van Wieringen, A., and Wouters, J. (2008). Apex 3: a multi-purpose test platform for auditory psychophysical experiments. *Journal of neuroscience methods*, 172(2):283–293.

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.

Hamilton, L. S., Edwards, E., and Chang, E. F. (2018). A spatial map of onset and sustained responses to speech in the human superior temporal gyrus. *Current Biology*, 28(12):1860–1871.

Harkrider, A. W., Plyler, P. N., and Hedrick, M. S. (2006). Effects of hearing loss and spectral shaping on identification and neural response patterns of stop-consonant stimuli. *The Journal of the Acoustical Society of America*, 120(2):915–925.

Harkrider, A. W., Plyler, P. N., and Hedrick, M. S. (2009). Effects of hearing loss and spectral shaping on identification and neural response patterns of stop-consonant stimuli in young adults. *Ear and hearing*, 30(1):31–42.

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

Koerner, T. K. and Zhang, Y. (2018). Differential effects of hearing impairment and age on electrophysiological and behavioral measures of speech in noise. *Hearing research*, 370:130–142.

Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., and Francart, T. (2019a). Predicting individual speech intelligibility from the cortical tracking of acoustic-and phonetic-level speech representations. *Hearing research*, 380:1–9.

Lesenfants, D., Vanthornhout, J., Verschueren, E., and Francart, T. (2019b). Data-driven spatial filtering for improved measurement of cortical tracking of multiple representations of speech. *Journal of neural engineering*, 16(6):066017.

Luts, H., Jansen, S., Dreschler, W., and Wouters, J. (2014). Development and normative data for the Flemish/Dutch matrix test.

Maamor, N. and Billings, C. J. (2017). Cortical signal-in-noise coding varies by noise type, signal-to-noise ratio, age, and hearing status. *Neuroscience letters*, 636:258–264.

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.

McClannahan, K. S., Backer, K. C., and Tremblay, K. L. (2019). Auditory evoked responses in older adults with normal hearing, untreated, and treated age-related hearing loss. *Ear and hearing*, 40(5):1106–1116.

Mirkovic, B., Debener, S., Schmidt, J., Jaeger, M., and Neher, T. (2019). Effects of directional sound processing and listener's motivation on eeg responses to continuous noisy speech: Do normal-hearing and aided hearing-impaired listeners differ? *Hearing Research*, 377:260–270.

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

Oates, P. A., Kurtzberg, D., and Stapells, D. R. (2002). Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. *Ear and hearing*, 23(5):399–415.

Oostenveld, R. and Praamstra, P. (2001). The five percent electrode system for high-resolution eeg and erp measurements. *Clinical neurophysiology*, 112(4):713–719.

Peelle, J. E. and Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends in neurosciences*, 39(7):486–497.

Petersen, E. B., Wöstmann, M., Obleser, J., and Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of neurophysiology*, 117(1):18–27.

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.

Presacco, A., Simon, J. Z., and Anderson, S. (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PloS one*, 14(3):e0213899.

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

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.

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.

Van Dun, B., Kania, A., and Dillon, H. (2016). Cortical auditory evoked potentials in (un) aided normal-hearing and hearing-impaired adults. In *Seminars in hearing*, volume 37, page 9. Thieme Medical Publishers.

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., and Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *Journal of the Association for Research in Otolaryngology*, 19(2):181–191.

Verschueren, E., Vanthornhout, J., and Francart, T. (2020). The effect of stimulus intensity on neural envelope tracking. *bioRxiv*.

Voeten, C. C. (2020). *buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression*. R package version 1.6.