Articles in PresS. J Neurophysiol (October 5, 2016). doi:10.1152/jn.00527.2016

Neural tracking of attended versus ignored speech is differentially affected by hearing loss

Eline Borch Petersen 1,2,4*,
Malte Wöstmann 3,
Jonas Obleser 3,
Thomas Lunner 1,2,4

1 Eriksholm Research Centre, Snekkersten, Denmark

2 Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

3 Department of Psychology, University of Lübeck, Lübeck, Germany

4 Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

* Correspondence: Eline Borch Petersen, Eriksholm Research Centre, Roertangvej 20, 3070 Snekkersten, Denmark. ebpe@eriksholm.com

Keywords: hearing loss, neural tracking, attention, speech-onset envelope, electroencephalography, cross-correlation

Running title: Neural speech tracking and hearing loss

Abstract

Hearing loss manifests as a reduced ability to understand speech, particularly in multi-talker situations. In these situations, younger normal-hearing listeners' brains are known to track attended speech through phase-locking of neural activity to the slow-varying envelope of the speech. This study investigates how hearing loss, compensated by hearing aids, affects the neural tracking of the speech-onset envelope in elderly participants with varying degree of hearing loss (*N* = 27, 62–86 years, hearing thresholds 11–73 dB hearing level). In an active listening task, a to-be-attended audiobook (signal) was either presented in quiet or against a competing to-be-ignored audiobook (noise), presented at three individualized signal-to-noise ratios (SNR). The neural tracking of the to-be-attended and to-be-ignored speech was quantified through the cross-correlation of the electroencephalogram (EEG) and the temporal envelope of speech. We primarily investigated the effects of hearing loss and SNR on the neural envelope tracking. First, we found that elderly hearing-impaired listeners' neural responses reliably track the envelope of to-be-attended speech more than to-be-ignored speech. Second, hearing loss relates to the neural tracking of to-be-ignored speech, resulting in a weaker differential neural tracking of to-be-attended versus to-be-ignored speech in listeners with worse hearing. Third, neural tracking of to-be-attended speech increased with decreasing background noise. Critically, the beneficial effect of reduced noise on neural speech tracking decreased with stronger hearing loss. In sum, our results show that a common sensorineural processing deficit, i.e., hearing loss, interacts with central attention mechanisms and reduces the differential tracking of attended and ignored speech.

New & Noteworthy

The current study investigates the effect of hearing loss in older listeners on the neural speech tracking of competing speech. Interestingly, we observe that whereas internal degradation (hearing loss) relates to the neural tracking of ignored speech, external sound degradation (ratio between attended and ignored speech; SNR) relates to tracking of attended speech. This provides the first evidence for hearing loss affecting the ability to neurally tracking speech.

Introduction

The ability to successfully distinguish between multiple talkers and selectively direct attention towards a particular speech stream is the heart of human communication (Cherry, 1953; McDermott, 2009). In such multi-talker situations, the neural response in the magneto/electroencephalogram (M/EEG) has been shown to phase-lock to the slow-amplitude fluctuations, often referred to as the broad-band “envelope”, of the speech signal. Neural phase-locking has been observed not only for speech, but for a variety of intelligible and unintelligible auditory stimuli (for review, see Ding and Simon, 2014; Zoefel and VanRullen, 2015). It has been proposed that upon neural detection of linguistic features, speech-specific brain regions are activated and higher-order processing initiated (Zoefel and VanRullen, 2015). As such, neural phase-locking is not solely driven by changes in the acoustic cue of the auditory stimuli, but also reflects cortical encoding and processing of the auditory signal. The phase-locking of neural activity to speech is often referred to as “neural tracking of speech” (Wöstmann et al., under revision; Zoefel and VanRullen, 2015). Interestingly, in a multi-talker situation, selective attention to one speaker results in stronger neural phase-locking to attended than ignored speech in younger normal-hearing listeners (Kerlin et al., 2010; Ding and Simon, 2012a; Mesgarani and Chang, 2012; Horton et al., 2013; O’Sullivan et al., 2015). This neural evidence for the processing of attended and ignored speech as separate auditory streams (Simon, 2015), supports previous behavioral studies showing that based on features from the auditory scene, attention can be exerted as to focus on a particular objects, while keeping other objects in the perceptual background (for review, see Shinn-Cunningham, 2008; Shinn-Cunningham and Best, 2008). This ability to perform attentional selection is essential for higher-level processing, such as successfully understanding the meaning of speech (Ding and Simon, 2014). Currently, the ability to perform neural speech tracking has only been investigated for younger normal-hearing listeners. Although it is known that listeners suffering from hearing loss (HL) experience great difficulties in multi-talker situation (Bronkhorst, 2000; Shinn-Cunningham and Best, 2008), it is unknown whether the deteriorating effect of HL on the afferent auditory signal cause changes in the neural tracking of speech.

Sensorineural HL causes distortion in the representation of auditory signals on the level of the cochlear, we say that HL causes an *internal* degradation of sounds. HL is often treated with hearing aids, through which incoming sounds are amplified in order to improve the audibility. Hence, hearing aids reduce the internal degradation and consequently relieve the cognitive resources deployed in correcting for the degradation on a higher-processing level (Pichora-Fuller et al., 1995; Lunner et al., 2009). However, despite adequate hearing-aid compensation, HL still affects the central processing of sound through reduced temporal precision (Tremblay and Ross, 2007) and contribute to gray-matter loss in the primary auditory cortex (Peelle et al., 2011). Behavioral studies also found that despite hearing-aid compensation, listeners with a HL experience deficits in the ability to: (1) process temporal fine-structure (Hopkins et al., 2008; Lunner et al., 2012), (2) understand speech in noise (Lunner, 2003), and (3) take advantage of spatial separation between talkers (Neher et al., 2009). The current study focuses on the possible effect of HL on the neural tracking of attended and ignored speech after hearing-aid compensation. This approach offers an alternative, and in some cases more realistic, way of looking at the effects of acoustic degradation.

Until now, the effect of acoustic degradation on neural speech tracking in younger normal-hearing listeners has only been investigated by *externally* degrading sounds. Since HL cause deficits in the ability to understand speech in noise and process temporal fine-structure, these two mechanisms of externally degrading the auditory signal presented to normal-hearing listeners are of special interest. Manipulating the signal-to-noise ratio (SNR) between the attended and ignored talker has been found to affect the neural tracking of attended speech around 50 ms after stimulus presentation (Ding and Simon, 2013a, SNR range: quiet, +6 to -9 dB). However, others have reported no effect of varying the SNR on the neural tracking of attended speech (Ding and Simon, 2012a, SNR range +8 to -8 dB; Kong et al., 2014, SNR range: quiet, +6, and 0 dB). More consistent findings have been reported on the effect of externally degrading the temporal fine-structure. Applying noise-vocoding of attended speech has been found to reduce the neural tracking of it when presented in quiet (Ding et al., 2013; Peelle et al., 2013), competing speech (Kong et al., 2015), and stationary noise (Ding et al., 2013). So far, studies have focused on the effect of external degradation of speech, while it is still unknown whether *internal* degradation of the auditory input, through sensorineural HL, influences the neural tracking of attended and ignored speech.

The aim of this study is twofold. First, we investigate how HL in elderly participants affects the neural tracking of attended and ignored speech in the EEG. Second, we test whether HL modulates the neural tracking of speech when altering the SNR between the attended and ignored talker. We hypothesize that listeners with more severe HL (i.e., with more internal degradation of sound) will exhibit reduced tracking of speech, evidenced by a diminished cross-correlation of the speech-onset envelope and the EEG response. Furthermore, we expect to find that lower SNRs (higher external sound degradation) result in stronger encoding of the competing ignored speech and weaker encoding of the attended speech relative to conditions with higher SNRs.

Methods

PARTICIPANTS

Twenty-seven native Swedish speaking participants (16 females, age range: 62–86 years) were recruited from the audiology clinic at the University Hospital of Linköping. The data from two additional participants were recorded, but excluded from all analyses due to a high degree of noise contamination in the EEG. All participants gave their written informed consent and the study was approved by the regional ethical board in Linköping, Sweden. For further details on participants' and methods, specifically the individualization of SNR levels, quantifying the HL, and recording of the EEG, see (Petersen et al., 2015).

Hearing abilities: Individual pure-tone audiometric thresholds for all participants are shown in **Figure 1A**. To obtain a single score reflecting the individuals' hearing ability, we calculated the pure-tone average (PTA) across the frequencies 0.5, 1, 2, 4, and 8 kHz, which ranged from 11 dB hearing level (normal hearing) to 73 dB hearing level (severe HL). The PTA was found to significantly increase with age ($r_{Pearson} = 0.398$, $p = 0.033$; **Figure 1B**).

EXPERIMENTAL DESIGN

Listening task: In the experiment, all participants were wearing Oticon Agil hearing aids (Oticon A/S, Smørum, Denmark) with individual quasi-linear amplification. In the current experiment, the dynamics of the auditory stimuli is slow-varying, hence the speech envelope is preserved through means of slow compression times (Dillon, 2001). All stimuli were presented directly through the hearing aids using the direct audio input (DAI), i.e., no free field presentation (see **Figure 1D**). The noise reduction algorithm and volume control of the hearing aids were turned off during the experiment. The experiment was conducted in an electrically shielded soundproof booth.

A 12-minutes section of the Swedish version of the audiobook 'Simple Genius' by David Baldacci, narrated by a male target talker (fundamental frequency 113.5 Hz), was presented diotically. In four intervals of three minutes each, the story was either narrated in quiet or masked at three different individualized SNR levels (see section on SNR individualization below; **Figure 1D**). The presentation order of the four SNR levels were randomized. The masker was a diotically presented female talker (fundamental frequency 179.5 Hz) narrating 'The Wonderful Adventures of Nils' by Selma Lagerlöf. Participants were instructed to attend to the male talker while ignoring the female talker. The duration of the pauses were computed for both speech signals and the distributions tested against each other to establish whether one talker had significantly longer pauses than the other. A two-sample Kolmogorov-Smirnov test revealed no difference in the pause-durations between the talkers (D(145) = 0.131, $p = 0.153$).

At the end of the listening task, participants were prompted with four question regarding the content of the attended story. The question were presented visually in a three-alternative forced-choice manner, with one question relating to the story heard during each of the four SNR levels.

<<Place Figure 1 here>>

SNR individualization: To avoid unequal intelligibility of the auditory stimuli due to differences in participants' hearing, individualized SNR levels were determined prior to the EEG experiment. The individualized SNR levels were estimated using the Swedish version of the hearing in noise test (HINT; Hällgren et al, 2006). In the HINT test, participants were presented with 40 spoken sentences embedded in speech-shaped steady-state noise at an output presentation level of 70 dB SPL, presented through the DAI of the hearing aids and amplified according to the individuals' audiogram. Using an adaptive tracking procedure (Levitt, 1971), the background noise level (measured as the SNR) at which each participant was able to repeat 80% of the words in a sentence was determined. This individual noise values is known as the Speech Reception Threshold (SRT) at 80% (SRT80). In the EEG experiment, the individual SRT80 level was used as the intermediate background-noise level for the participant (denoted 0 dB SRT80). By raising and lowering the SNR by 4 dB from the 0 dB SRT80 level, the more favorable (+4 dB SRT80) and less favorable (–4 dB SRT80) SNR levels were created. As such, a listener with an SRT80-value of -1 dB SNR was subjected to background-noise levels at +3 dB SNR (+4 dB SRT80), -1 dB SNR (0 dB SRT80), and –5 dB SNR (–4 dB SRT80). Of the 81 recorded conditions (27 participants and 3 background-noise levels) where attended and ignored speech were presented simultaneously, 20 of them (24.7%) had SNRs at or below 0 dB. Practically, both the level of the to-be-attended and the to-be-ignored signal were adjusted to maintain a constant presentation level of 70 dB SPL, before hearing-aid amplification.

The individually determined SRT80 level had an average value of 4.61 dB (standard error of the mean (SEM) = 0.86 dB, range –1 to 12.7 dB). A significant increase in the SRT80 value with higher PTA was found ($r_{Pearson} = 0.768$, $p < 0.001$), i.e., participants with worse hearing required better SNR to maintain a performance of 80%.

EEG RECORDING AND ANALYSIS

Data recording and preprocessing: The EEG was recorded using the EGI system (Electrical Geodesic Inc., Eugene, OR, USA) with 103 scalp electrodes at a sampling frequency of 250 Hz. Offline, the raw EEG data were bandpass-filtered between 0.5 and 45 Hz using an 6th order Butterworth filter, and re-referenced from Cz to the mean of the left and right mastoids. All analyses were done using customized MATLAB scripts (R2011b, Mathworks Inc.) and the Fieldtrip toolbox (Oostenveld et al., 2011).

Independent component analysis (ICA) was performed on the continuous data and components corresponding to eye blinks, saccadic eye movements, muscle activity, and heartbeats were identified by visual inspection of components' topographies and time courses and rejected. The data were projected back to electrode-time space before the continuous recordings were separated into four 3-minute segments based on the SNR level applied to the particular segment.

Calculation of neural speech tracking: The speech-onset envelopes were extracted by first calculating the absolute of the Hilbert transform of the speech signals. This transform was low-pass filtered at 25 Hz (3rd order Butterworth filter) and the first derivative was taken before it was half-wave rectified and down sampled to the sampling frequency of the EEG (250 Hz) (Hambrook and Tata, 2014). By taking the first derivative of the speech envelope (hence denoted speech-onset envelope), the salient changes in the speech signal are emphasized, specifically at the onset of tones and syllables. Practically, using the first derivative of the speech envelop removes potential drift in the correlation between EEG and speech envelope.

For each of the four 3-minute segments, seventeen 10-second epochs were extracted from each channel of the EEG, disregarding the first and last 5 seconds of each segment. To measure how well the neural response phase-locked to the envelope of the speech stimuli, we used the cross-correlation. In detail, for each 10-second epoch and channel, three cross-correlations were calculated between the EEG and (1) the speech-onset envelope of the to-be-attended talker, (2) the speech-onset envelope of the to-be-ignored talker, and (3) the speech-onset envelope of the to-be-attended talker taken from a random part of the story (i.e., non-time-aligned) in order to obtain a control condition of the overall correspondence between the EEG signal and the speech-onset envelope. From hereon the three cross-correlations will be denoted the "attended", the "ignored", and the "control" condition, respectively (see **Figure 1D**).

In general, the cross-correlation measures the similarity between the EEG response and the speech-onset envelope as a function of temporal displacement between the two signals, i.e., time-lag. The cross-correlation coefficients ($r_{crosscorr}$) possibly range between –1 and +1, with values closer to 0 indicating no resemblance and values near ±1 indicating a perfect linear correspondence of the EEG response and speech-onset envelope. Whereas the cross-correlations with attended and ignored speech both reflect the encoding of speech being presented to the participants, the control condition takes into account the temporal characteristics of the attended talker, but without being systematically related to the particular segment of EEG it was correlated with.

The effect of attention on the neural tracking of speech was quantified by subtracting the cross-correlation coefficients of the ignored condition from that of the attended condition (i.e., attended–ignored) for each participant and SNR level. In one 3-minute segment of the listening task, the attended speech was presented in quiet; hence, no ignored response could be calculated. Consequently, the ignored and attended–ignored conditions included only responses for the three SNR levels were the competing talker was presented (+4, 0, and –4 dB SRT80).

STATISTICAL ANALYSES

Statistical effects of the categorical factor of SNR level (quiet, +4 dB SRT80, 0 dB SRT80, and –4 dB SRT80) experimentally varied within subjects and the continuous covariate HL (measured as rPTA, see below) varying between subjects, on the cross-correlations were investigated. Critically, for the investigation of the cross-correlation responses in the active listening conditions (attended, ignored, and attended-ignored), the control condition acted as a baseline by testing the remaining conditions against the control.

Statistical elimination of age-effects from the measure of hearing loss: In order to investigate the effect of HL on the neural tracking of speech, irrespective of possible effects of participants' age, we utilized the residuals resulting from the linear regression of PTA on age. The z-scored residualized PTA will be referred to as *rPTA* and employed in all further analysis (the same measure was used before by Petersen et al., 2015).

Behavioral data: Whether the proportion of correct answers differed between SNR levels were tested using a Chi-square test. The relationship between the accuracy and HL was investigated using Pearson's correlation between the proportion of correct answers pooled across SNR levels for each participant and rPTA.

Neural tracking of attended and ignored speech: Statistical comparisons between the control condition and the three active listening conditions were done using the cluster-based approach implemented in the Fieldtrip toolbox (Maris and Oostenveld, 2007). Dependent-samples t-tests between the control and each active listening condition for each time-lag (time resolution 0.004 s) and electrode were conducted. Based on the resulting t-values, clusters were formed by connecting adjacent time samples with p-values < 0.05 containing at least three neighboring electrodes. Within each cluster, the single-sample t-values were summed and compared to a permutation-distribution. The permutation-distribution consisted of summed t-values from clusters generated through 1000 iterations of randomly assigning time-electrode samples to one of the two compared conditions. The summed t-values of clusters derived from the condition-contrast of interest were compared with the summed t-values from the permuted clusters (Maris and Oostenveld, 2007). A cluster was considered significant if the sum of its t-values exceeded the 95%-percentile of the permutation distribution, corresponding to a one-sided p-value < 0.05. In the following, all cluster-based tests had setting as described above, unless otherwise stated.

Neural speech tracking as a function of SNR level: A two-step approach was used to investigate the effect of SNR level on the neural speech tracking. First, assuming that noise-induced changes in $r_{crosscorr}$ would be linearly related to the SNR level, cluster-based independent-samples regression analysis was used on the single-subject level. For each participant, the change of the $r_{crosscorr}$ in the attended and ignored conditions as a function of the three SNR levels was investigated by ranking the conditions; +4 dB SRT80, 0 dB SRT80, –4 dB SRT80 and assigning them the linearly-spaced contrast-coefficients –1, 0, and +1, respectively. The regression analysis implemented in the Fieldtrip toolbox, assumes equal separation between the independent variables (SNR level). This criteria is only fulfilled for the three SNR levels where ignored speech is presented (spaced by 4 dB), but not for the quiet condition (infinite SNR) which was not included in the statistical cluster-analysis. Second, the resulting linear regression coefficients across participants (β-weights; quantifying the linear change in $r_{crosscorr}$ with increasing SNR) were tested against zero using cluster-based dependent-samples t-tests on the group level.

Effects of hearing loss on neural speech tracking: Whether HL asserted an effect on the neural tracking of speech was investigated using Pearson’s correlation. From the time-lags and electrodes showing a significant difference in the tracking of attended and ignored speech, $r_{crosscorr}$-values were extracted for each participant and correlated with rPTA. Pearson’s correlation was also applied to investigate the interaction between HL and SNR level by correlating rPTA and the difference in speech tracking between SNR levels. For each participant, the difference in speech tracking was calculated by subtracting the average $r_{crosscorr}$-value within the significant cluster from the most favorable SNR level (quiet for the attended condition and +4 dB SRT80 for the ignored condition) from that of the least favorable SNR level (-4 dB SRT80 for both conditions).

Results

INTELLIGIBILITY ENSURED ACROSS SNR LEVELS

The performance accuracy (see **Figure 1C**), proved to be significantly higher than change level, lying at 33.33% for a three-alternative forced choice task ($\chi^2(1) = 29.67$, $p > 0.001$). No significant difference in the proportion of correct answers were found between SNR levels ($\chi^2(3) = 2.49$, $p = 0.48$). The Pearson's correlation between showed no relationship between the performance calculated across SNR levels and rPTA ($r_{Pearson} = 0.05$, $p = 0.81$).

OLDER LISTENERS NEURALLY TRACK ATTENDED MORE THAN IGNORED SPEECH

The cross-correlation coefficients ($r_{crosscorr}$) from the four conditions (attended, ignored, attended–ignored, and control) are shown in **Figure 2A**. As expected, the control condition exhibited values of $r_{crosscorr}$ close to zero across all time-lags (range $-2.5 \cdot 10^{-4}$ to $+2.5 \cdot 10^{-4}$). This indicates no systematic relationship between the EEG response and the speech-onset envelope presented in another time-lag interval. The $r_{crosscorr}$ of the attended and ignored conditions averaged across SNR levels ranged from –0.01 to +0.01.

For the neural tracking of attended speech, the cluster-based analysis identified three time intervals which differed significantly from the control condition (see **Figure 2A**; blue clusters): A significant positive deflection peaking at 75 ms (time-lag 24–104 ms, 74 electrodes, $p < 0.001$), a negative deflection peaking at 150 ms (time-lag 112–212 ms, 83 electrodes, $p < 0.001$), and a positive deflection peaking at 250 ms (time-lag 220–356 ms, 64 electrodes, $p < 0.001$). From hereon, these three deflections will be denoted $P1_{crosscorr}$, $N1_{crosscorr}$, and $P2_{crosscorr}$, respectively.

For the neural tracking of ignored speech, the statistical analysis revealed a significant $P1_{crosscorr}$ (time-lag 16–104 ms, 81 electrodes, $p < 0.001$) and $P2_{crosscorr}$ (time-lag 196–292 ms, 72 electrodes, $p = 0.002$) compared to the control condition (**Figure 2A**; red clusters). A cluster was identified around $N1_{crosscorr}$ for the ignored condition (time-lag 136–152 ms, 47 electrodes), however the summed t-values within the cluster only approached statistical significance ($p = 0.073$). Most importantly, the attentional modulation (i.e., attended–ignored) significantly differed from the control condition in the time-lag interval including $N1_{crosscorr}$ and $P2_{crosscorr}$ (time-lag 108–232 ms, 83 electrodes, $p < 0.001$, **Figure 2A**; black cluster), which indicates stronger neural tracking of attended than ignored speech within this time-lag interval.

<< Place Figure 2 here>>

ATTENTIONAL MODULATION OF SPEECH TRACKING DECREASES WITH HEARING LOSS

The linear effect of HL (rPTA) on the attentional modulation of neural speech tracking (attended–ignored condition) was investigated by extracting values of $r_{crosscorr}$ from the time-lags and electrodes where the attended–ignored condition differed significantly from the control (black cluster in **Figure 2A**). We found a significant decrease in the attentional modulation of neural speech tracking with worse hearing ($r_{Pearson}$ = 0.542, p = 0.004, **Figure 2B** left), indicating that listeners with stronger HL exhibit similar neural tracking of attended and ignored speech.

The significant relationship between HL and the individual SRT80-values ($r_{Pearson}$ = 0.751, p < 0.001) could suggest that the individualized SNR-levels, rather than HL, were affecting the attentional modulation. However, a multiple regression analysis (F(2,25) = 3.66, p = 0.027, R-squared adjusted = 0.235), revealed no significant effect of SRT80 (p = 0.834) or of the interaction between rPTA and SRT80 (p = 0.488) on attentional modulation. The only significant predictor of attentional modulation was hearing loss (rPTA, p = 0.012).

To test whether HL was associated with the tracking of attended or ignored speech, $r_{crosscorr}$-values from the time-lag and electrodes showing a significant attentional modulation (black cluster in **Figure 2A**), were extracted separately for the attended and ignored conditions separately and correlated with HL. Whereas the tracking of attended speech showed no significant relationship with HL ($r_{Pearson}$ = 0.096, p = 0.633), tracking of the ignored speech showed a significant linear decrease in magnitude with worse hearing ($r_{Pearson}$ = –0.515, p = 0.006, **Figure 2B** right). Visual inspection of the cross-correlation responses of the ignored talker (data not shown) revealed that participants with normal hearing had smaller $N1_{crosscorr}$-peaks and consequent earlier $P2_{crosscorr}$-peaks, compared to participants with worse hearing. Consequently, this resulted in more positive $r_{crosscorr}$-values for tracking of the ignored talker within the attentional modulation cluster for participants with better hearing. This indicates that participants with worse hearing are unable to suppress the ignored talker, resulting in higher similarity in the neural tracking of attended and ignored speech, evident from the declining attentional modulation.

EXTERNAL NOISE REDUCES THE NEURAL TRACKING OF ATTENDED SPEECH

Figure 3A shows cross-correlations between the EEG response and the envelope of attended speech for the three different SNR levels where ignored speech was presented (+4 dB SRT80, 0 dB SRT80, and –4 dB

SRT80). Two significant clusters were identified in which $r_{crosscorr}$ of attended speech significantly varied with SNR level: A cluster in the time-lag interval of the $N1_{crosscorr}$ (denoted C1, time-lag 124–160 ms, 72 electrodes, *p* = 0.006) and a cluster in the time-lag interval of the $P2_{crosscorr}$ (denoted C2, time-lag 228–268 ms, 55 electrodes, *p* = 0.028). The $r_{crosscorr}$-values extracted from C1 and C2 for each SNR level revealed that tracking of the attended speech increased in magnitude with lower noise levels within both clusters (**Figure 3B**). Although not included in the statistical analysis, the quiet condition showed a further increase in neural tracking of attended speech (grey bars in **Figure 3B**). For the sake of comparison, the $r_{crosscorr}$-values for the ignored-speech tracking within C1 and C2 are plotted in red in **Figure 3B**. Note that the high $r_{crosscorr}$-values for ignored condition within C2 is caused by an earlier peak in $P2_{crosscorr}$ compared to the encoding of the attended speech (see **Figure 2A**).

A cluster-based statistical test found no significant effect of SNR level on the neural tracking of ignored speech (all *ps* > 0.36).

<< Place Figure 3 here>>

HEARING LOSS MODULATES TRACKING OF ATTENDED SPEECH AT DIFFERENT SNR LEVELS

We investigated the interaction between HL and SNR level by utilizing the difference in neural tracking between the most and least favorable SNR level. **Figure 3B** shows that the quiet condition, although not included in the statistical analysis, supported the finding that less background noise resulted in better neural tracking of attended speech. Therefore, the quiet condition was included into the computation of the $r_{crosscorr}$-difference for the attended speech (quiet minus –4 dB SRT80). **Figure 4A** shows the $r_{crosscorr}$-difference for each individual sorted according to the degree of HL (rPTA), for the two clusters C1 and C2 (identified in **Figure 3A**). Pearson's correlations revealed a significant decrease in the $r_{crosscorr}$-difference (quiet minus –4 dB SRT80; blue lines in **Figure 4A**) with worse hearing for the C1 cluster ($r_{Pearson} = 0.394$, $p = 0.042$), with the $r_{crosscorr}$-differences from the C2 cluster suggesting a similar trend ($r_{Pearson} = –0.349$, $p = 0.075$). In other words, in the neural tracking of attended speech, participants with better hearing showed a larger sensitivity to changes in the SNR level. Participants with worse hearing show no change in the tracking of the attended speech between the least favorable SNR level (–4 dB SRT80) and the quiet condition, see individual data in **Figure 4B**.

As expected, the $r_{crosscorr}$-difference for the ignored talker, calculated between the SNR levels +4 dB SRT80 and –4 dB SRT80 (+4 dB SRT80 minus –4 dB SRT80), showed no significant relationship with rPTA within the C1 and C2 clusters (both *p*s > 0.13, red lines in **Figure 4A**).

<< Place Figure 4 here>>

Discussion

The present study used a competing-talker paradigm to investigate the neural response to continuous speech in elderly listeners with varying degrees of hearing loss (HL) and under varying degrees of signal-to-noise (SNR) levels. We asked how both factors, internal HL and external SNR degradation, would interfere with the neural tracking of speech. Our results can be summarized as follows: (i) Older listeners' with varying degree of HL reliably track the speech-onset envelope of attended speech, more than that ignored speech. (ii) Worse hearing relates to reduced attentional modulation in the neural speech tracking, driven by a higher similarity in the tracking of attended and ignored speech. (iii) More favorable SNR in the acoustic stimulation improves the neural tracking of attended speech, but this improvement diminishes with more severe HL.

ATTENTION MODULATES SPEECH TRACKING IN ELDERLY LISTENERS WITH VARYING DEGREE OF HEARING LOSS

In line with recent findings for younger normal-hearing listeners, three significant components ($P1_{crosscorr}$, $N1_{crosscorr}$, $P2_{crosscorr}$) were identified in the neural tracking of attended speech for our older listeners with varying degrees of HL (see **Figure 2A**; Power et al., 2012; Horton et al., 2013; Kong et al., 2014; O'Sullivan et al., 2015). Peaks in the neural speech tracking response are thought to reflect different processing stages, from the encoding of auditory features ($P1_{crosscorr}$) to evaluating the behavioral importance of the auditory object ($N1_{crosscorr}$ and $P2_{crosscorr}$; Ding and Simon, 2013b). Although not identified in all previous studies, we observed significant $P2_{crosscorr}$-components for both the attended and ignored condition. Horton and colleagues suggest that the emergence of the $P2_{crosscorr}$ depends on the difficulty of the experimental task (Horton et al., 2013). Horton and colleagues also observed a change in polarity for $N1_{crosscorr}$ suggestive of an enhancement of the attended and suppression of the ignored, respectively, for younger normal-hearing listeners. No change in the $N1_{crosscorr}$-polarity was observed in the current study, which might suggest that attentional modulation was more difficult to assert in the current study than in the study by Horton and colleagues. The general compliance in cross-correlation magnitude and response pattern between this study and previous studies in younger listeners, suggests that also elderly subjects with varying degrees of HL exhibit reliable neural tracking of speech.

Previous studies have found attention to modulate speech tracking around 150 ms ($N1_{crosscorr}$) within the neural speech tracking of normal-hearing younger listeners (Ding and Simon, 2012a, 2012b; Power et al., 2012; Hambrook and Tata, 2014; Kong et al., 2014). Interestingly, the cluster-based approach in the current study allowing for a more detailed analysis, revealed attentional modulation not only of $N1_{crosscorr}$, but of the $N1_{crosscorr}$–$P2_{crosscorr}$ complex. Since ageing, like hearing impairment, is associated with a decline in the ability to assert attentional control (Pichora-Fuller and Singh, 2006; Passow et al., 2012) profound age-effects on the attentional modulation of neural speech tracking might be expected. However, the observed significant difference between the neural tracking of attended and ignored speech suggests that attentional modulation is asserted in the neural response of older listeners.

HEARING LOSS REDUCES THE ATTENTIONAL MODULATION OF NEURAL SPEECH TRACKING

In line with our hypothesis, HL had a detrimental effect on the attentional modulation of neural speech tracking (**Figure 2B**). Specifically, we observed that hearing loss was associated with changes in the tracking of ignored speech, rather than tracking of attended speech. In other words, participants with worse hearing showed a higher similarity in the neural tracking of attended and ignored speech. This suggests that HL deteriorates the segregation of competing talkers, resulting in deficient inhibition of the ignored speech signal. This might explain why listeners suffering from HL report difficulties in coping with multi-talker situation, even when wearing hearing aids (Bronkhorst, 2000; Shinn-Cunningham and Best, 2008).

As the individualized background-noise levels result in mainly positive SNRs, it could be speculated that the neural tracking of attended speech was favored as its relative level in the speech mixture exceeds that of the ignored speech. Indeed significantly, higher SNRs (SRT80) were applied for participants with worse hearing, which could potentially cause the observed attentional modulation effect. However, as we observed no significant relationship between attentional modulation and the individualized background-noise levels (SRT80), we do not suspect the application of positive SNRs to affect the attentional modulation. It must be emphasized that although worse hearing is associated with significantly higher SRT80-values, poorer cognitive abilities are also known to reduce the ability to understand speech in noise, thus influencing the SRT80-value irrespective of hearing loss (Lunner, 2003; Petersen et al., 2016).

From a cognitive perspective, internal degradation (HL) poses additional constraints on the limited cognitive resources involved in listening, leaving fewer resources for the perceptual processing of the auditory input (Pichora-Fuller et al., 1995; Lunner et al., 2009). Research on ageing has established that particularly the ability to inhibit irrelevant information is reduced with age (Hasher and Zacks, 1988; Hasher et al., 2008). Hasher and Zacks (1988) note that deficits in the inhibitory process allow irrelevant information to disrupt the selective-attention process and thereby occupy cognitive resources. Our findings suggests that worse hearing, like increased age, affects the ability to inhibit irrelevant information, evident from the increased neural tracking of ignored speech.

It is well-established that HL is associated with difficulties in processing temporal fine-structure (Hopkins et al., 2008; Lunner et al., 2012), hence parallels can be drawn between HL and the effect of vocoding the speech material presented to normal-hearing listeners (Shannon et al., 2007). Indeed, reducing temporal fine-structure in a competing-talker task has been found to induce a decline in attentional modulation in younger normal-hearing listeners, resulting from changes in the tracking of both the attended and ignored speech (Kong et al., 2015). Our result showed no effect of HL on attended speech tracking possibly resulting from HL causing other processing deficiencies than just a reduced sensitivity to temporal fine-structure (Moore, 2007).

BACKGROUND NOISE REDUCES THE NEURAL TRACKING OF ATTENDED SPEECH

Effects of increasing the background-noise level (by decreasing the SNR from +4 dB to –4 dB SRT80) on the tracking of attended speech were found within two time-electrode clusters, both showing values of $r_{crosscorr}$ closer to zero with higher levels of background noise (i.e., lower SNRs, **Figure 3**). This finding supports part of our hypothesis that lower SNRs result in weaker tracking of attended speech. Hence, since the cluster-based analysis showed no effect of SNR on the tracking of ignored speech, the hypothesis that tracking of the ignored speech would increase with lower SNRs is not supported. Generally, external degradation of speech is not always found to affect the neural speech tracking (e.g., see Howard and Poeppel, 2010). Also studies specifically altering the SNRs between talkers do not always show an effect on the neural speech tracking (Ding and Simon, 2012a; Kong et al., 2014). However, it must be considered that the elderly participants with varying degree of hearing loss could apply another listening strategy in multi-talker situations.

While the sparse behavioral measure showed no effect SNR level, we suspect that low number of questions asked for each participant causes this non-significant effect of background noise level. However, the behavioral data shows that participants were performing above chance level, suggesting that the attended speech was intelligible (see **Figure 1C**). We therefore do not suspect that the detrimental effect of SNR level on the neural tracking of attended speech to be caused by an unintelligible stimuli. Indeed, we have previously found task performance to be high (>80%), but modulated by the background noise level in an auditory Sternberg task when using the same individualized noise levels and the same participants as included in the current study (Petersen et al., 2015).

Although the statistical approaches used to identify effects of internal and external sound degradation differ, it is interesting to note that we found HL and SNR to be associated with the neural representation of ignored and attended speech, respectively. It has previously been suggested that the neural representations of attended and ignored speech are neurally processed independently, on the level of separate auditory objects (Simon, 2015). Following this line of argumentation, it is possible for internal auditory degradation (HL) and external sound degradation (SNR) to affect the two auditory objects (attended and ignored speech) independently. When the SNR of attended relative to ignored speech was increased, we observed that the neural representation of attended speech was enhanced, while the neural representation of ignored speech was unaffected. Since a larger part of the neural tracking response for attended speech differs from zero, compared to the response to ignored speech (**Figure 2A**), this increases the likelihood of observing SNR level effects on the tracking of attended than ignored speech.

However, why does HL have a stronger impact on the neural tracking of the ignored speech? An enhanced neural tracking response for a particular speech stream at a time-lag of ~150ms (around $N1_{crosscorr}$) reflects attentional modulation, manifesting as a deeper encoding of the attended speech stream rather than the ignored (see **Figure 2A**; Ding and Simon, 2012a, 2012b; Power et al., 2012; Hambrook and Tata, 2014; Kong et al., 2014). HL reduces the spectro-temporal dissimilarity between attended and ignored speech already on the level of the cochlea (Moore, 2007), which impairs the formation of separate auditory objects for the two speech signals (for review, see Shinn-Cunningham and Best, 2008). Consequently, listeners with more severe HL show a deep encoding of the attended, but also the ignored speech signal. In other words, our results suggest that listeners with more severe HL track the entire auditory scene (attended and ignored speech) without neurally inhibiting the ignored speech. This could relate to the difficulties experienced by hearing-impaired listeners' in complex multi-talker situations (Shinn-Cunningham and Best, 2008).

Considering the experimental design, differences in the neural tracking of the attended and ignored speech could be affected by the difference in the speech characteristics of the two talkers. Previous studies have found no significant effects of gender on the neural speech tracking response in younger normal-hearing listener during active listening (Ding and Simon, 2012a; Kong et al., 2015). Although, we would not expect that age and hearing loss would cause an interaction between neural tracking and talker characteristics, we are not able to test this claim with the current experimental design.

HEARING LOSS REDUCES SENSITIVITY TO CHANGING NOISE LEVELS

Analyzing the change in the neural tracking of attended speech between the quiet and least favorable SNR level (–4 dB SRT80) revealed that participants with worse hearing did not improve the speech tracking as the SNR improved (**Figure 4**). As such, participants with worse hearing seem insensitive to changes in the SNR level, contrary to participants with better hearing, who show a larger difference between the tracking of the attended talker in quiet and at -4 dB SRT80.

A similar effect of HL on the sensitivity to noise has been observed in the pupil response of older listeners (Zekveld et al., 2011). Zekveld and colleagues argue that speech information processing is more superficial for listeners with HL, in that they perform less information storage and semantic processing, which leads to reduced pupil responses, as a measure of listening effort, in the older participants with HL. The interaction between HL and SNR observed in the present study suggests that the insensitivity to changes in the SNR level could result in superficial speech information processing, proposed by Zekveld and colleagues. Interestingly, a recent study showed that the EEG response tracks not only the speech envelope of natural speech, but also the phonetic and spectral features important for higher-level processing and understanding of speech (Di Liberto et al., 2015). In relation to HL, a link between neural speech tracking and higher-level processing, could explain why hearing-impaired listeners have problems not only understanding speech in noise, but also in coding of information into the long-term memory (Rönnberg et al., 2011).

In summary, our results demonstrate that older participants with varying degrees of hearing loss under aided listening conditions show surprisingly robust neural tracking of speech. Furthermore, the internal degradation through the loss of hearing results in reduced attentional modulation of neural speech tracking, mainly driven by limited inhibition of ignored speech. Interestingly, manipulating external degradation, by lowering the SNR, manifests in a reduced ability to neurally track attended speech. Participants with worse hearing showed no improvement in attended speech tracking with lowered background noise.

Thus, internal and external sound degradation affect different aspects of auditory speech processing, either by reducing inhibition of ignored speech (internal degradation) or reducing neural encoding of the attended speech (external degradation). In addition, hearing-aid amplification in itself is seen not to restore normal neural tracking of the auditory input for participants suffering from a hearing loss. This corroborates the sustained difficulties in everyday multi-talker situations often reported by listeners suffering from hearing loss.

Acknowledgements

EBP, TL, and JO are supported the Oticon Foundation. JO and MW are supported by an ERC Consolidator grant to JO (ERC-Cog-2014 AUDADAPT), and JO and TL are supported by the Volkswagen foundation. We wish to thank the participants in this study, as well as Gunilla Wänström, Irene Slättengren, Mathias Hällgren, and Stefan Stenfelt for their assistance during the experiment.

Disclosures

Eriksholm Research Centre (EBP, TL) is part of Oticon A/S, Smørum, Denmark.

References

Bronkhorst AW. The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple Talker Conditions. *Acta Acust united with Acust* 86: 117–128, 2000.

Cherry EC. Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25: 975–979, 1953.

Dillon H. Hearing aids. 1st ed. Thieme.

Ding N, **Chatterjee M**, **Simon JZ**. Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage* 88: 41–46, 2013.

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

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

Ding N, **Simon JZ**. Robust Cortical Encoding of Slow Temporal Modulations of Speech. In: *Basic Aspects of Hearing: Physiology and Perception*, edited by Moore BCJ. Springer Science+Business Media New York, p. 373–381.

Ding N, **Simon JZ**. Cortical entrainment to continuous speech: Functional roles and interpretations. *Front Hum Neurosci* 8: 1–7, 2014.

Hambrook DA, **Tata MS**. Theta-band phase tracking in the two-talker problem. *Brain Lang* 135: 52–56, 2014.

Hasher L, **Lustig C**, **Zacks R**. Inhibitory Mechanisms and the Control of Attention. In: *Variation in Working Memory*, edited by Conway A, Jarrold C, Kane M, Miyake A, Towse J. Oxford University Press, p. 227–249.

Hasher L, **Zacks RT**. Working memory, comprehension, and aging: A review and new view. *Psychol Learn Motiv* 22: 193–225, 1988.

Hopkins K, **Moore BCJ**, **Stone MA**. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J Acoust Soc Am* 123: 1140–1153, 2008.

Horton C, **D'Zmura M**, **Srinivasan R**. Suppression of competing speech through entrainment of cortical oscillations. *J Neurophysiol* 109: 3082–3093, 2013.

Howard MF, **Poeppel D**. Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *J Neurophysiol* 104: 2500–11, 2010.

Hällgren M, **Larsby B**, **Arlinger S**. A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *Int J Audiol* 45: 227–37, 2006.

Kerlin JR, **Shahin AJ**, **Miller LM**. Attentional gain control of ongoing cortical speech representations in a "cocktail party". *J Neurosci* 30: 620–628, 2010.

Kong YY, **Mullangi A**, **Ding N**. Differential modulation of auditory responses to attended and unattended speech in different listening conditions. *Hear Res* 316: 73–81, 2014.

Kong Y-Y, **Somarowthu A**, **Ding N**. Effects of Spectral Degradation on Attentional Modulation of Cortical Auditory Responses to Continuous Speech. *J Assoc Res Otolaryngol* 16: 783–796, 2015.

Levitt H. Transformed Up-Down Methods in Psychoacoustics. *J Acoust Soc Am* 49: 467–477, 1971.

Di Liberto GM, **O'Sullivan JA**, **Lalor EC**. Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Curr Biol* 25: 2457–2465, 2015.

Lunner T. Cognitive function in relation to hearing aid use. *International J Audiol* 42: S49–S58, 2003.

Lunner T, **Hietkamp RK**, **Andersen MR**, **Hopkins K**, **Moore BCJ**. Effect of Speech Material on the Benefit of Temporal Fine Structure Information in Speech for Young Normal-Hearing and Older Hearing-Impaired Participants. *Ear Hear* 33: 377–388, 2012.

Lunner T, **Rudner M**, **Rönnberg J**. Cognition and hearing aids. *Scand J Psychol* 50: 395–403, 2009.

Maris E, **Oostenveld R**. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164: 177–190, 2007.

McDermott JH. The Cocktail Party Problem. *Curr Biol* 19: R1024–R1027, 2009.

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

Moore BCJ. Cochlear Hearing Loss. 2nd editio. John Wiley and Sons, Ltd.

Neher T, **Behrens T**, **Carlile S**, **Jin C**, **Kragelund L**, **Petersen AS**, **Schaik A Van**. Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *Int J Audiol* 48: 758–774, 2009.

O'Sullivan JA, **Power AJ**, **Mesgarani N**, **Rajaram S**, **Foxe JJ**, **Shinn-Cunningham BG**, **Slaney M**, **Shamma S a**, **Lalor EC**. Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cereb cortex* 25: 1697–1706, 2015.

Oostenveld R, **Fries P**, **Maris E**, **Schoffelen J-M**. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011: 156869, 2011.

Passow S, **Westerhausen R**, **Wartenburger I**, **Hugdahl K**, **Heekeren HR**, **Lindenberger U**, **Li S-C**. Human aging compromises attentional control of auditory perception. *Psychol Aging* 27: 99–105, 2012.

Peelle JE, **Gross J**, **Davis MH**. Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cereb Cortex* 23: 1378–1387, 2013.

Peelle JE, **Troiani V**, **Grossman M**, **Wingfield A**. Hearing loss in older adults affects neural systems supporting speech comprehension. *J Neurosci* 31: 12638–12643, 2011.

Petersen EB, **Lunner T**, **Vestergaard MD**, **Sundewall Thorén E**. Danish reading span data from 283 hearing-aid users, including a sub-group analysis of their relationship to speech-in-noise performance. *Int J Audiol* 55: 1–8, 2016.

Petersen EB, **Wöstmann M**, **Obleser J**, **Stenfelt S**, **Lunner T**. Hearing loss impacts neural alpha oscillations under adverse listening conditions. *Front Audit Cogn Neurosci* 6: 1–11, 2015.

Pichora-Fuller MK, **Schneider BA**, **Daneman M**. How younger and old adults listen to and remember speech in noise. *J Acoust Soc Am* 97: 593–608, 1995.

Pichora-Fuller MK, **Singh G**. Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif* 10: 29–59, 2006.

Power AJ, **Foxe JJ**, **Forde E-J**, **Reilly RB**, **Lalor EC**. At what time is the cocktail party? A late locus of selective attention to natural speech. *Eur J Neurosci* 35: 1497–1503, 2012.

Rönnberg J, **Danielsson H**, **Rudner M**, **Arlinger S**, **Sternäng O**, **Wahlin Å**, **Nilsson L-G**. Hearing Loss is Negatively Related to Episodic and Semantic Long-Term Memory but Not to Short-Term Memory. *J Speech, Lang Hear Res* 54: 705–727, 2011.

Shannon R V, **Zeng F**, **Kamath V**, **Wygonski J**, **Ekelid M**. Speech Recognition with Primarily Temporal Cues Robert. *Science (80-)* 270: 303–304, 2007.

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

Shinn-Cunningham BG, **Best V**. Selective attention in normal and impaired hearing. *Trends Amplif* 12: 283–299, 2008.

Simon JZ. The Encoding of Auditory Objects in Auditory Cortex: Insights from Magnetoencephalography. *Int J Psychophysiol* 95: 184–190, 2015.

Tremblay K, **Ross B**. Effects of age and age-related hearing loss on the brain. *J Commun Disord* 40: 305–312, 2007.

Wöstmann M, **Fiedler L**, **Obleser J**. Tracking the signal, cracking the code: Speech and speech comprehension in non-invasive human electrophysiology. *Languange, Cogn Neurosci* in review, under review.

Zekveld AA, **Kramer SE**, **Festen JM**. Cognitive Load During Speech Perception in Noise : The Influence of Age , Hearing Loss , and Cognition on the Pupil Response. *Ear Hear* 32: 498–510, 2011.

Zoefel B, **VanRullen R**. The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound. *Front Hum Neurosci* 9: 651, 2015.

Figure captions

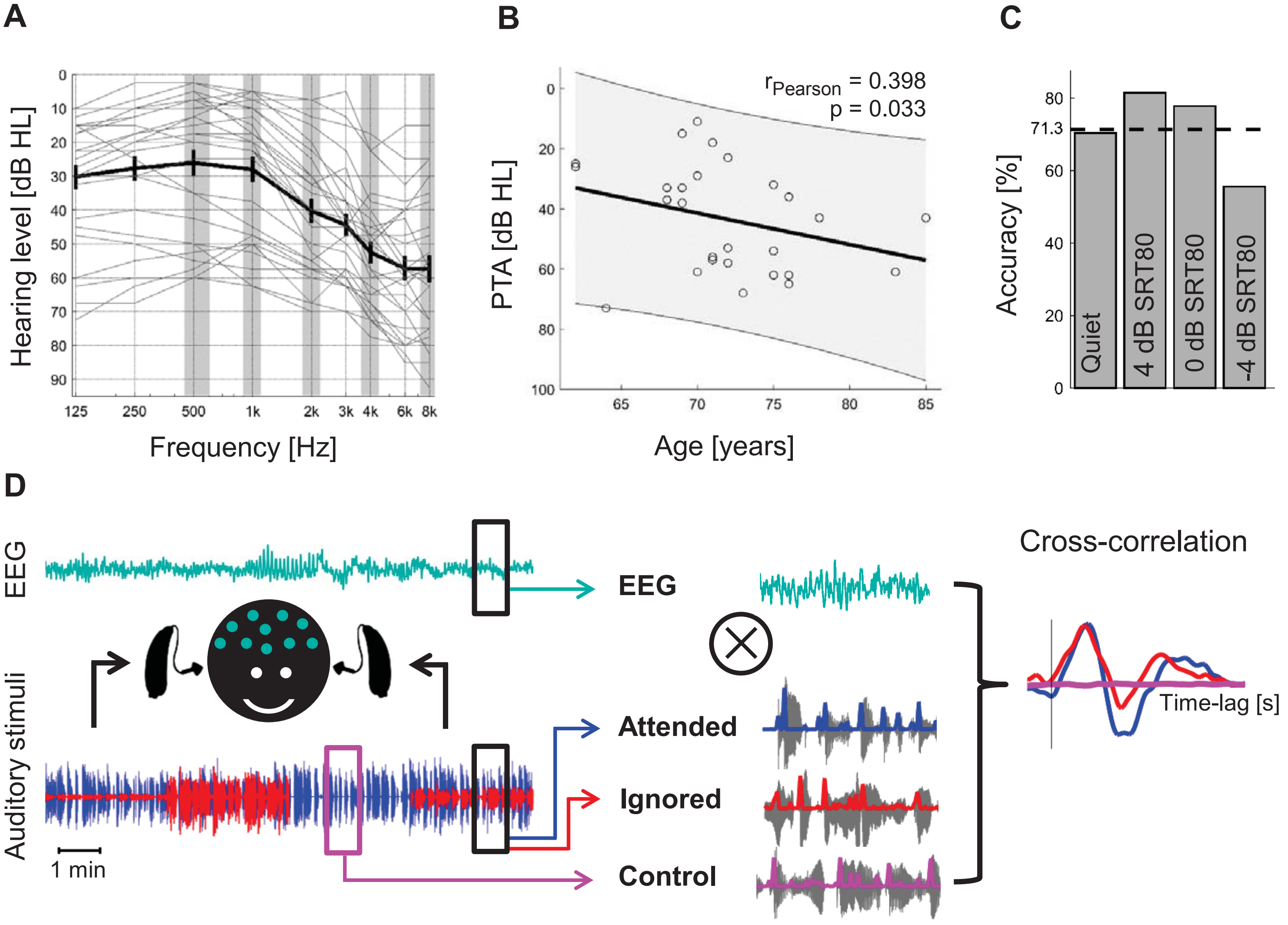
Figure 1: **Hearing abilities and experimental design. (A)** Pure-tone hearing thresholds for the each participant averaged between ears are shown in thin grey lines. The average hearing threshold across all subjects is shown in black with error bars indicating ± 1 SEM. The pure-tone average (PTA) was calculated as the average across the frequencies highlighted with gray shading. **(B)** The significant linear decrease in hearing ability (quantified as PTA) with age ($p = 0.033$) is shown with the lease-square regression line (bold black line). The 95% confidence interval of the regression is indicated with thin lines. **(C)** Response accuracy for the questions regarding the content of the attended story for the four SNR levels. The percentage of correct answers is calculated across participants. The average accuracy across SNR levels is 71.30% (dashed line). **(D)** Left, bottom: Outline of the acoustic stimuli; a to-be-attended audiobook (male talker, blue) and a to-be-ignored audiobook (female talker, red). The to-be-attended talker was presented in quiet or masked by the to-be-ignored talker at three SNR levels. Left, top: All sounds were presented to both ears through hearing aids. The scalp EEG (illustrated with cyan dots and lines) was recorded during the task. Right: To quantify the neural tracking of speech, the broad-band speech-onset envelope of the to-be-attended (blue line) and to-be-ignored (red line) speech signals were extracted and cross-correlated with the EEG response (cyan) for all electrodes. For statistical analysis, a control condition was created by correlating the EEG response with a randomly picked, non-time-aligned, segment of the to-be-attended talker (magenta).

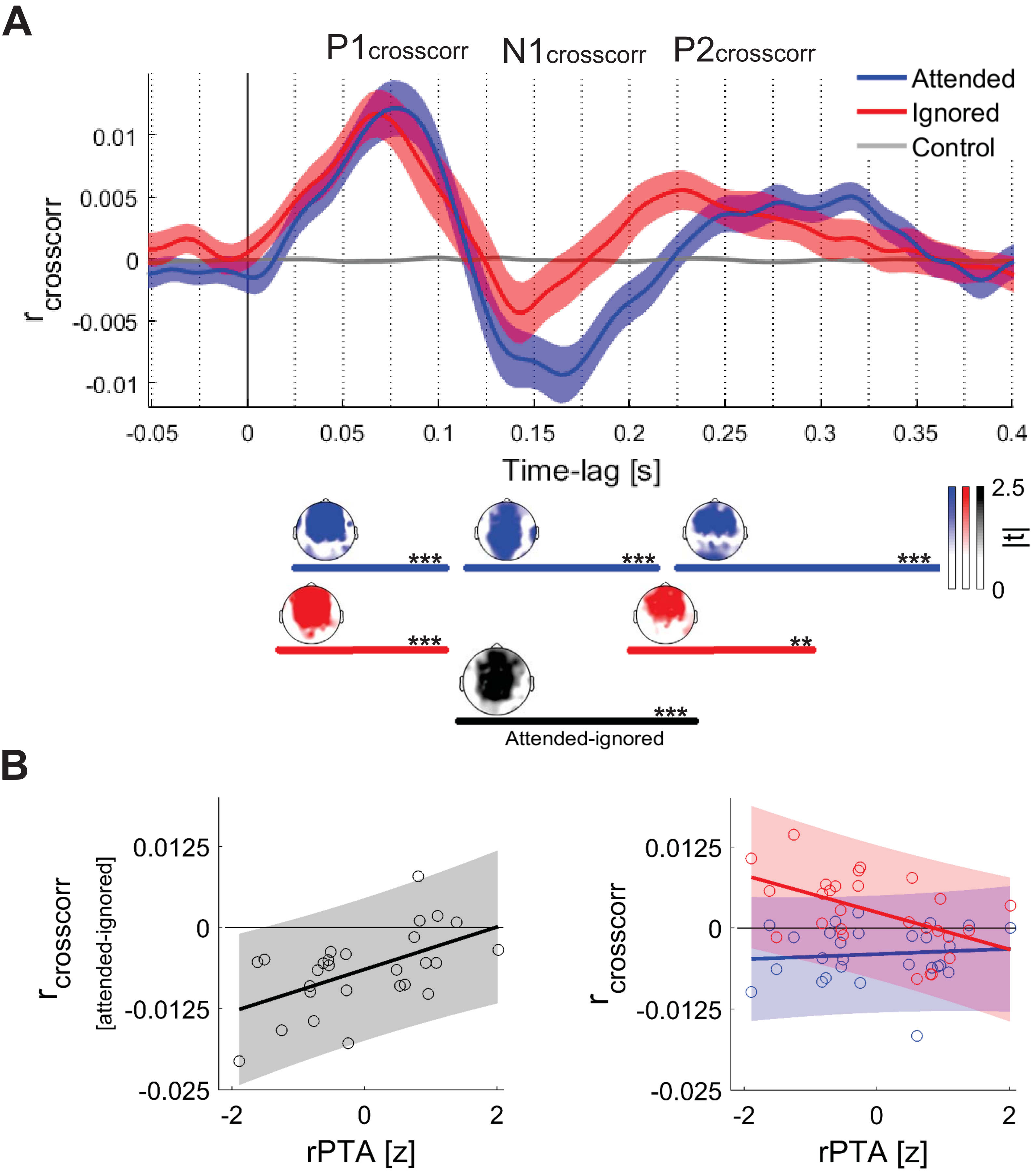
Figure 2: Neural tracking of speech-onset envelopes and effect of hearing loss. (A) Top: Solid lines and shaded areas respectively show the grand-average cross-correlation (across $N = 27$ participants, the 58 electrodes common for all significant clusters, and all SNR levels) and the 95% confidence intervals for attended speech (blue), ignored speech (red), and the control condition (grey). Notation of the three components $P1_{\text{crosscorr}}$, $N1_{\text{crosscorr}}$, and $P2_{\text{crosscorr}}$ is shown above the responses. Bottom: Results of the cluster-based permutation tests (see text for details). Time-lags at which the active listening conditions differ significantly from the control condition are indicated with horizontal bars (blue, attended speech; red, ignored speech; black, attended–ignored). The corresponding topographic maps of the t-values are positioned above the bars. Asterisks indicate the p-values for each cluster (*** $p < 0.001$, ** $p < 0.01$). **(B)** Left: The significant linear least-squares regression between hearing loss (rPTA) and the attentional modulation (attended–ignored, $p = 0.004$) extracted from the significant attended–ignored cluster (black in Figure 2A). Right: From the significant time-lags and electrodes of the attended–ignored cluster, values of $r_{\text{crosscorr}}$ for the ignored condition (red, $p = 0.006$), but not for the attended condition (blue, $p = 0.633$),

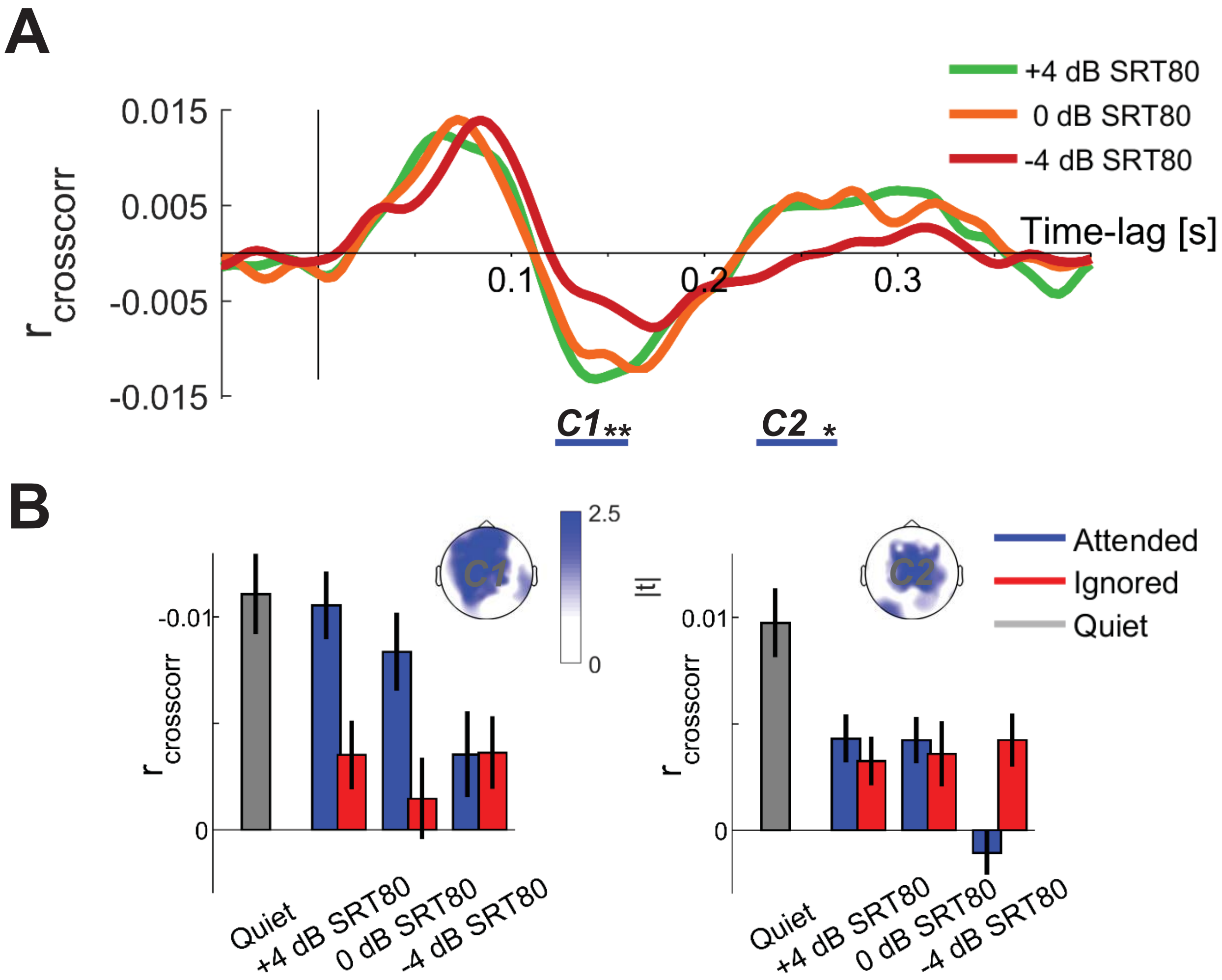
significantly correlated with hearing loss (rPTA). The shaded areas indicate the 95% confidence interval of the regression lines.

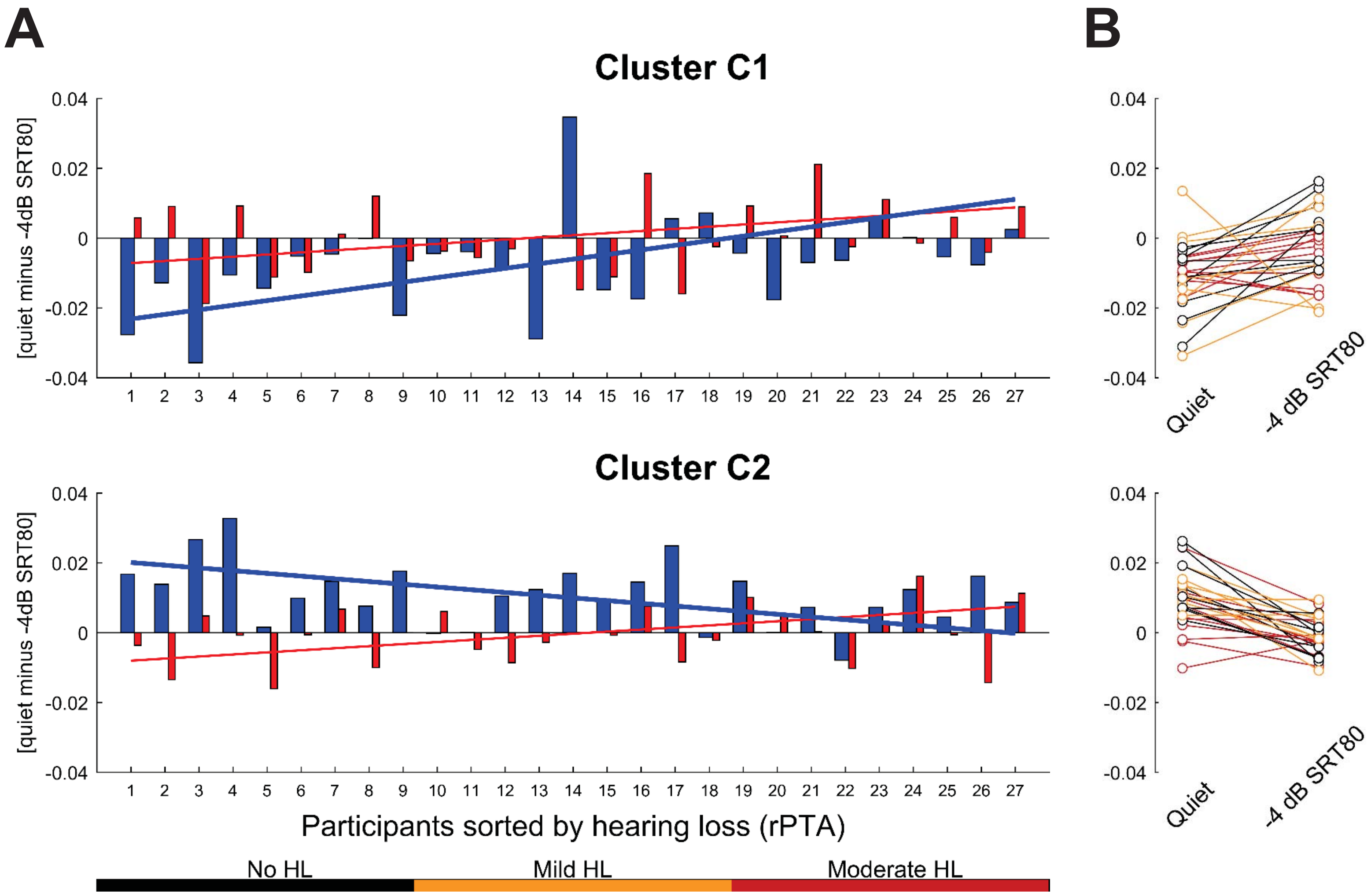
Figure 3: Effects of SNR level on the neural tracking of attended speech. (A) Solid lines show the grand-average cross-correlations for attended speech (across *N* = 27 participants and the 44 electrodes common for both significant clusters) for the three SNR levels where ignored speech was presented (green, +4 dB SRT80; orange, 0 dB SRT80; red, –4 dB SRT80). Horizontal blue bars show the temporal extent of the two significant clusters (denoted C1 and C2) exhibiting a linear effect of SNR level on the tracking of attended speech. Asterisks indicate the p-values for each cluster (** $p < 0.01$, * $p < 0.05$). **(B)** Topographic maps show the spatial extend of the two significant clusters (C1 on the left, C2 on the right, note that the y-axes are reversed). The averaged $r_{crosscorr}$-values from the significant time-lags and electrodes are shown for tracking of attended (blue) and ignored (red) speech for the three SNR levels where ignored speech was presented. For comparison, the tracking of attended speech during the quiet condition is also shown (grey, not included in the statistical analysis). Error bars indicate ±1 SEM.

Figure 4: Interaction between SNR level and hearing loss on the tracking of attended speech. (A) Data for each participant, ordered according to the degree of hearing loss (rPTA), is presented in bars. Individual differences in $r_{crosscorr}$ between the quiet and the –4 dB SRT80 condition for tracking of attended speech within the two significant clusters identified in Figure 3A (in blue, top left: C1, bottom left: C2). For comparison, the tracking of ignored speech, calculated as the difference in $r_{crosscorr}$ between the 4 dB SRT80 and the –4 dB SRT80 condition within the two clusters, are shown in red bars. The linear least-squares regressions between HL and the $r_{crosscorr}$-differences are shown in solid lines for the attended speech (blue, C1: $p = 0.042$, C2: $p = 0.075$) and ignored speech (red, C1: $p = 0.223$, C2: $p = 0.13$). **(B)** Individual $r_{crosscorr}$-value for tracking of the attended speech for the quiet and -4 dB SRT80 condition from C1 (top) and C2 (bottom). The individual lines are color-coded according to hearing loss, by separating the participants into three groups of equal size (n = 9; black, no hearing loss; orange, mild hearing loss; red, moderate hearing loss).


A
Hearing level [dB HL]
0
10
20
30
40
50
60
70
80
90
125
250
500
1k
2k
3k
4k
6k
8k
Frequency [Hz]
B
PTA [dB HL]
$r_{Pearson} = 0.398$
$p = 0.033$
0
20
40
60
80
100
65
70
75
80
85
Age [years]
C
Accuracy [%]
0
20
40
60
71.3
80
Quiet
4 dB SRT80
0 dB SRT80
-4 dB SRT80
D
EEG
Auditory stimuli
1 min
EEG
Attended
Ignored
Control
Cross-correlation
Time-lag [s]


A
P1crosscorr
N1crosscorr
P2crosscorr
Attended
Ignored
Control
0.01
0.005
0
-0.005
-0.01
$r_{crosscorr}$
-0.05
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
Time-lag [s]
***
***
***
***
**
***
Attended-ignored
2.5
0
|t|
B
0.0125
0
-0.0125
-0.025
$r_{crosscorr}$ [attended-ignored]
-2
0
2
rPTA [z]
0.0125
0
-0.0125
-0.025
$r_{crosscorr}$
-2
0
2
rPTA [z]


A
+4 dB SRT80
0 dB SRT80
-4 dB SRT80
0.015
0.005
-0.005
-0.015
$r_{crosscorr}$
0.1
0.2
0.3
Time-lag [s]
C1**
C2 *
B
2.5
0
|t|
C1
C2
-0.01
0
$r_{crosscorr}$
Quiet
+4 dB SRT80
0 dB SRT80
-4 dB SRT80
0.01
0
$r_{crosscorr}$
Quiet
+4 dB SRT80
0 dB SRT80
-4 dB SRT80
Attended
Ignored
Quiet


A
B
Cluster C1
Cluster C2
r_crosscorr
[quiet minus -4dB SRT80]
0.04
0.02
0
-0.02
-0.04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Participants sorted by hearing loss (rPTA)
No HL
Mild HL
Moderate HL
Quiet
-4 dB SRT80