



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2022

---

**Better speech-in-noise comprehension is associated with enhanced neural speech tracking in older adults with hearing impairment**

Schmitt, Raffael ; Meyer, Martin ; Giroud, Nathalie

DOI: <https://doi.org/10.1016/j.cortex.2022.02.017>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-225913>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Schmitt, Raffael; Meyer, Martin; Giroud, Nathalie (2022). Better speech-in-noise comprehension is associated with enhanced neural speech tracking in older adults with hearing impairment. *Cortex*, 151:133-146.

DOI: <https://doi.org/10.1016/j.cortex.2022.02.017>



## Research Report

# Better speech-in-noise comprehension is associated with enhanced neural speech tracking in older adults with hearing impairment



Raffael Schmitt <sup>a,\*</sup>, Martin Meyer <sup>b,c,d,e,f</sup> and Nathalie Giroud <sup>a,c</sup>

<sup>a</sup> Department of Computational Linguistics, University of Zurich, Zurich, Switzerland

<sup>b</sup> Department of Comparative Language Science, University of Zurich, Zurich, Switzerland

<sup>c</sup> Center for Neuroscience Zurich, University and ETH of Zurich, Zurich, Switzerland

<sup>d</sup> University Research Priority Program (URPP) Dynamics of Healthy Aging, University of Zurich, Zurich, Switzerland

<sup>e</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Zurich, Switzerland

<sup>f</sup> Cognitive Psychology Unit, Alpen-Adria University, Klagenfurt, Austria

## ARTICLE INFO

## Article history:

Received 20 August 2021

Reviewed 22 October 2021

Revised 19 December 2021

Accepted 3 February 2022

Action editor Alessandro Tavano

Published online 19 March 2022

## Keywords:

Neural speech tracking

Age-related hearing loss

EEG

Speech comprehension

## ABSTRACT

The alignment between low-frequency activity in the brain and slow acoustic modulations in the speech signal depicts a core principle in present theories of speech perception—a process referred to as ‘neural speech tracking’. While most older adults, particularly those with highly prevalent age-related hearing loss, have difficulties with speech perception and comprehension, the impact of hearing loss on neural speech tracking is still unclear. In this study we investigated the effects of pure-tone hearing loss and different types of background noise on the neural tracking response in a large sample of older adults ( $N = 101$ ). Furthermore, we examined whether the neural tracking response was predictive for speech comprehension. For this purpose, we obtained scalp EEG from our participants who had varying degrees of pure-tone hearing loss (7.5–59.6 dB HL for .5–8 kHz pure tones) while they listened to sentences in quiet, pink and multi-talker babble noise. Speech tracking was quantified by computing the cross-correlation between the EEG signal and the amplitude envelope of the sentences heard. A higher degree of pure-tone hearing loss was associated with greater neural speech tracking (i.e., greater cross-correlation). Additionally, neural speech tracking showed a positive association with speech comprehension. This relationship was modulated by the degree of pure-tone hearing loss with hearing-impaired participants benefitting more from greater neural speech tracking. Our results highlight the potential of neural speech tracking as an objective measure of speech comprehension and as a possible target mechanism for clinical interventions such as neurofeedback. Furthermore, the interaction between speech tracking and pure-tone hearing loss suggests

\* Corresponding author. University of Zurich, Department of Computational Linguistics, Computational Neuroscience of Speech & Hearing, Andreasstrasse 15, CH-8050 Zürich, Switzerland.

E-mail addresses: [raffael.schmitt@uzh.ch](mailto:raffael.schmitt@uzh.ch) (R. Schmitt), [martin.meyer@uzh.ch](mailto:martin.meyer@uzh.ch) (M. Meyer), [nathalie.giroud@uzh.ch](mailto:nathalie.giroud@uzh.ch) (N. Giroud).

<https://doi.org/10.1016/j.cortex.2022.02.017>

0010-9452/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

a compensatory mechanism by which the hearing-impaired rely more on slow amplitude modulations in the speech signal.

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

By the age of 65, roughly 33% of men and 29% of women experience a hearing loss of 35 dB or more (Homans et al., 2017). The significance of these figures becomes clear when one considers the possible long-term consequences an impaired hearing can have: studies suggest that hearing loss is associated with social isolation (e.g., Ciorba, Bianchini, Pelucchi, & Pastore, 2012), depression (e.g., Lawrence et al., 2020), cognitive decline (e.g., Lin et al., 2013) and incident dementia (e.g., Loughrey, Kelly, Kelley, Brennan, & Lawlor, 2018). Thus, hearing loss in the aging population constitutes a major public health issue (World Health Organization, 2017).

The disruption of understanding speech in everyday communication is characteristic for age-related hearing loss, especially in the presence of interfering background noise (e.g., Humes et al., 2012). Besides the age-related decline of cochlear structures, alterations along the auditory pathway up to the cortex contribute to impaired speech perception faced by hearing-impaired older adults (Gates & Mills, 2005). In other words, the integrity of the inner ear (as often measured by pure-tone audiometry) explains only parts of these speech perception and comprehension problems. However, the contributions of age-related changes in the brainstem and cortex are far from being understood (Peelle & Wingfield, 2016). Therefore, studying pure-tone hearing loss and its interaction with neural processing of speech is crucial for understanding age-related hearing loss and its accompanying perceptual deficits.

When listening to a continuous stream of speech, low-frequency activity in the auditory cortex of young normal hearing listeners is known to synchronize to temporal regularities (i.e., the amplitude envelope) in the speech signal (Luo & Poeppel, 2007). A prominent hypothesis suggests that this ‘neural speech tracking’ enables the segmentation of continuous speech into discrete chunks of acoustic units, ultimately building the foundation for subsequent decoding steps, whereat these acoustic units are mapped onto linguistic units in the mental dictionary (Ghitza, 2013; Giraud & Poeppel, 2012; Poeppel & Assaneo, 2020; for other functional hypotheses proposed see; Ding & Simon, 2014). In other words, neural tracking responses to a continuous speech stream are not solely driven by passively elicited responses, but rather reflect cortical encoding and processing of higher-level acoustic and linguistic speech features. Importantly, accumulating evidence from studies using concurrent electrical stimulation attribute a causal role to neural speech tracking for spoken language comprehension (Riecke, Formisano, Sorger, Başkent, & Gaudrain, 2018; Wilsch, Neuling, Obleser, & Herrmann, 2018; Zoefel, Archer-Boyd, & Davis, 2018). It is therefore unsurprising that research of recent years has shown an increasing interest in neural synchronization to continuous speech when

investigating the detrimental effects of hearing loss in older adults.

This line of research has shown that aging is associated with enhanced neural speech tracking responses (Decruy, Vanthornhout, & Francart, 2019; Presacco, Simon, & Anderson, 2016). The relationship between pure-tone hearing loss and neural speech tracking on the other hand is far from clear. Presacco, Simon, and Anderson (2019) compared hearing-impaired and normal hearing older adults as well as a normal hearing young control group and found augmented neural responses to speech envelope modulations in the older adults. However, the driving factor seemed to be aging as there was no difference between the two older groups. In a study by Petersen, Wöstmann, Obleser, and Lunner (2017), hearing loss was associated with enhanced processing of an ignored speech stream while the processing of the target stream was unaffected. The authors interpret this weakened differential processing of target and ignored speech as a possible cause of problems that the hearing-impaired experience in multi-talker situations. Studies by Decruy, Vanthornhout, and Francart (2020) as well as Gillis, Decruy, Vanthornhout, and Francart (2021) on the other hand showed that envelope tracking was amplified in hearing-impaired listeners compared to an age-matched normal hearing control group. This is in line with Fuglsang, Märcher-Rørsted, Dau, and Hjortkjær (2020) who reported enhanced envelope-entrained cortical responses for a target speech stream in a group of hearing-impaired older adults compared to normal hearing age-matched peers. Lastly, the study by Mirkovic, Debener, Schmidt, Jaeger, and Neher (2019) corroborates these results (Decruy et al., 2020; Fuglsang et al., 2020; Gillis et al., 2021) where greater envelope tracking of a target talker was observed in participants with hearing loss (although no statistics were reported).

Based on these studies, it seems plausible that the enhanced response in hearing-impaired adults reflects a compensatory mechanism where the brain becomes more susceptible to the slow varying temporal envelope of speech. Notwithstanding, this assumption would receive significantly stronger support if it were to be shown that speech tracking is of greater importance for perceptual processes (e.g., speech intelligibility or comprehension) in the hearing-impaired compared to normal hearing controls. Studies addressing this question by relating speech-brain alignments with direct measures of speech understanding in hearing-impaired adults are, however, scarce and limited to small sample sizes (Decruy et al., 2020; Gillis et al., 2021; Kurthen et al., 2021).

It is also worth noting that in some of the above-mentioned studies (Decruy et al., 2020; Fuglsang et al., 2020; Gillis et al., 2021; Mirkovic et al., 2019; Petersen et al., 2017) hearing was corrected by hearing aids. This might have

obscured the true relationship between hearing loss and the neural tracking response as compression in hearing aids is known to alter the amplitude envelope (Souza, 2002). Even if this was taken into account by providing hearing-impaired participants with linear amplification during the experiment, the possibility of lasting perceptual effects that a long-term usage of a hearing aid might have as well as potential confounds of low adaptation to linear amplifications in this short time period of usage could not be ruled out (Giroud, Lemke, Reich, Matthes, & Meyer, 2017). We therefore aimed to close this gap by considering participants with uncorrected hearing loss.

Given the inconsistent findings regarding the relationship between pure-tone hearing loss and speech-brain synchronization, a central aim of the current study was to investigate this association in a large group of older adults. We analysed data from 101 older adults with varying degrees of uncorrected pure-tone hearing loss (i.e., pure-tone average (PTA) over .5–8 kHz between 7.5 dB HL and 59.6 dB HL) and measured their neural activity while they listened to continuous speech in quiet and in two different types of noise (i.e., pink (1/f noise) and multi-talker babble noise).

Lastly, we investigated whether neural speech tracking benefits perceptual performance in an intelligibility and a comprehension task and if this relationship alters as a function of hearing sensitivity. A similar approach has been adopted in previous studies (Decruy et al., 2020; Gillis et al., 2021; Kurthen et al., 2021) which found a positive association between neural speech tracking and different measures of speech understanding albeit this relationship being the same in both normal hearing and hearing-impaired participants. If, however, speech tracking was indeed amplified by hearing loss, and if this additional amplification represented compensation, we expected the relationship between tracking and behavioural measures of speech perception to be modulated by the degree of hearing loss.

## 2. Materials and methods

### 2.1. Subjects

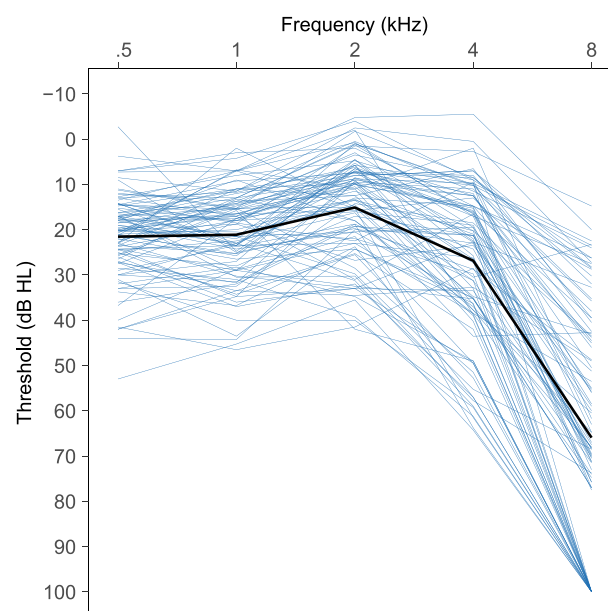
A sample of 101 right-handed older adults ( $M_{\text{Age}} = 70.52$ ,  $\text{range}_{\text{Age}} = 63\text{--}80$ ,  $SD_{\text{Age}} = 3.99$ , 47 female) participated in the present study. Participants showed no cognitive impairment (Montreal Cognitive Assessment  $>26$ ; Nasreddine et al., 2005) nor neurological or psychiatric disorders. Their native language was Swiss German and they had not learned a second language before the age of seven. Participants reported no signs of speech or language deficits (i.e., dyslexia) and professional musicians were excluded. Participants' hearing loss did not exceed 60 dB for the octave frequencies from .5 to 8 kHz ( $M_{\text{PTA}} = 30.1$ ,  $\text{range}_{\text{PTA}} = 7.5\text{--}59.6$ ,  $SD_{\text{PTA}} = 11.64$ ) and pure-tone thresholds were nearly symmetrical for both ears ( $<15$  dB interaural threshold difference). All participants gave their written informed consent and received monetary compensation for participation. The study was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee (Cantonal Ethics Committee Zurich, application no. 2017-00284).

### 2.2. Audiometry

Pure-tone audiometric thresholds were measured using a MATLAB-based probe-detection paradigm thoroughly described in previous studies (Giroud et al., 2018; Lecluyse & Meddis, 2009; Lecluyse, Tan, McFerran, & Meddis, 2013). The PTAs for each participant were then calculated by averaging thresholds across the octave frequencies from .5 to 8 kHz. Pure-tone hearing thresholds for each participant are shown in Fig. 1. During the audiometric assessment and the subsequent experimental task, participants were seated in an electrically shielded soundproof booth. Stimuli were controlled via sound card (RME Babyface Pro, RME, Haimhausen, Germany) and presented through a loudspeaker with linear frequency response (8030B Studio Monitor, Genelec, Iisalmi, Finland).

### 2.3. Stimuli and experimental paradigm

Participants listened to sentences that were spoken in standard German by a trained female speaker at a mean speech rate of 4.66 syllables/s (range = 4.03–5.61). The speech rate was estimated automatically in Praat (version 6.1.40; Boersma & Weenink, 2021) using the algorithm of de Jong and Wempe (2009) for syllable nuclei detection. To ensure a neutral content that would activate as little prior knowledge as possible, sentences contained information about EU regulations. Sentences were either presented in quiet (“Quiet” condition) or masked by pink (1/f) (“Pink” condition) or multi-talker babble noise (“Babble” condition) while the presentation order of the conditions was randomized. Babble noise consisted of eight other, randomly selected overlaid sentences while silent pauses were cut out to make “hearing in the gaps” impossible.



**Fig. 1 – Pure-tone audiometry.** Hearing thresholds (dB HL) for frequencies from .5 to 8 kHz. Each line represents an individual participant. The mean hearing threshold is depicted as black line.

In each condition, 30 sentences with an average length of 10.34 sec (range = 8.41–12.35) were presented at 70 dB SPL, ensuring audibility for all participants. The signal-to-noise ratio (SNR) for conditions Pink and Babble was set to 0 which was based on a pilot study that revealed the best performance distribution for that SNR level. Noise faded in after 1.5 sec of target sentence onset which enabled the participants to follow the target signal. This was especially required in the Babble condition, as background noise was made up of the same speaker and following the target signal would have been nearly impossible.

After each sentence an intelligibility task was conducted. Namely, a 300 msec sound snippet was presented whereupon participants had to indicate whether the snippet was part of the previously presented sentence or not. Behavioural response was given by either clicking the left (yes) or right (no) mouse button. The snippets were randomly cut out from the last 3 to .3 sec of each sentence to ensure minimal working memory effort. Additionally, after every fifth sentence a four-alternative forced-choice comprehension question was displayed on the screen. Participants responded by clicking the numbers “1”, “2”, “3” or “4” on the keyboard, indicating which of the four answers to the comprehension question was correct. The experiment started with a training session in which the participants performed the above-mentioned intelligibility and comprehension tasks in quiet. After every trial, a feedback was displayed on screen and the training trial was repeated until the participant performed without mistake to ensure that the task was understood correctly.

#### 2.4. EEG recording and preprocessing

EEG was continuously recorded with 128 Ag/AgCl scalp electrodes (BioSemi ActiveTwo, BioSemi, Amsterdam, the Netherlands) and digitized at a sampling rate of 512 Hz. The data were online bandpass filtered between .1 and 100 Hz, while impedances were kept below 25 k $\Omega$ . Preprocessing was performed in MATLAB (version R2020a, Mathworks) using the FieldTrip toolbox (version 20180128; Oostenveld, Fries, Maris, & Schoffelen, 2011). Offline, the EEG data were segmented into trials covering the respective sentence with a baseline of 2 sec. For preprocessing, data were re-referenced to Cz and a bandstop-filter was applied for 49–51 Hz (twopass Butterworth, 3rd order, Hamming window) in order to eliminate artifacts resulting from electrical interference. EEG data were further bandpass filtered between .1 and 30 Hz (twopass Butterworth, 3rd order, Hamming window) and visually scanned for bad channels, which were removed. Subsequently, data were re-referenced to the average reference and an independent component analysis was applied (Jung et al., 2000) while components corresponding to eye blinks and heart activity were removed. Identification was done based on visual inspection of the component's time course and topography. Noisy channels were then interpolated using spherical spline interpolation. Afterwards, data were resampled to 128 Hz, bandpass filtered between 2 and 8 Hz (twopass Butterworth, 3rd order, Hamming window) and baseline corrected. Finally, EEG data were cut to a uniform length of 5 sec (i.e., from 3 to 8 sec) to minimize sentence and noise onset as well as accompanying evoked responses.

#### 2.5. Speech envelope extraction

Based on the analyses of Biesmans, Das, Francart, and Bertrand (2017), speech envelopes were extracted using a gammatone filter bank followed by power law, mimicking the compressive response of the inner ear. More precisely, the raw acoustic signals were passed through a filter bank containing 24 bandpass filters with an equivalent rectangular bandwidth of 1 and centre frequencies ranging from 100 Hz to 4 kHz. The output from each filter was full-wave rectified and power-law compressed (i.e., taking the absolute value and raising it to the power of .6). The resulting sub-band envelopes were then averaged, culminating in a single envelope which was further resampled to 128 Hz in order to match the sampling rate of the EEG data. In a final step, envelopes were bandpass filtered between 2 and 8 Hz (twopass Butterworth, 3rd order, Hamming window) and cut from 3 to 8 sec to match the bandwidth and length of the EEG signal. A passband of 2–8 Hz was selected as the auditory system appears to be optimally tuned to temporal modulations in this frequency range (e.g., Doelling, Arnal, Ghitza, & Poeppel, 2014; Teng, Tian, Rowland, & Poeppel, 2017; for a review see; Poeppel & Assaneo, 2020). Envelope extraction was performed in MATLAB (version R2020a, MathWorks) using the Auditory Modeling Toolbox (Søndergaard & Majdak, 2013).

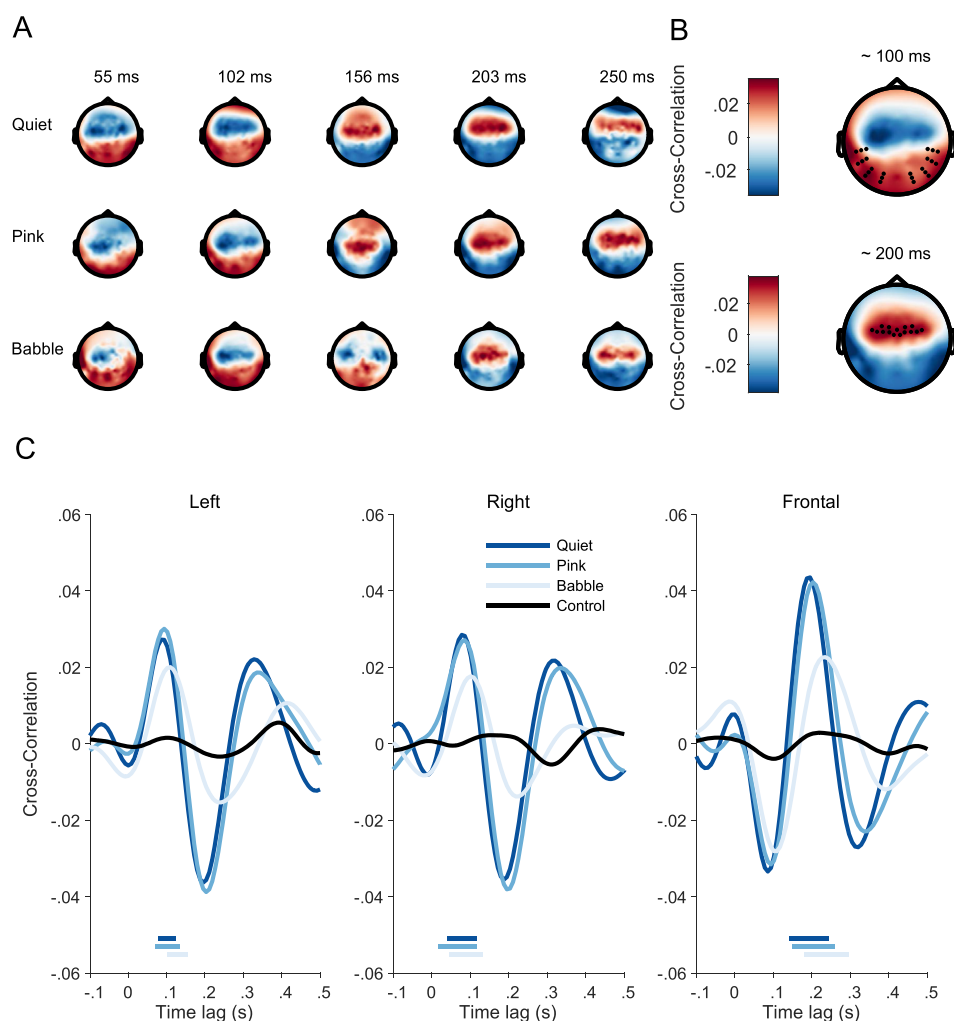
#### 2.6. Neural speech tracking

Neural tracking of speech was quantified by means of cross-correlation which estimates the similarity between two time series (e.g., speech envelope and EEG) as a function of temporal displacement (e.g., Horton, D'Zmura, & Srinivasan, 2013). The resulting coefficients range from  $-1$  to  $+1$ , where a value of  $+1$  (resp.  $-1$ ) denotes a completely positive (resp. negative) linear relationship between the two time series. In our study, each EEG electrode was cross-correlated with the envelope of the presented sentence with a temporal resolution of approximately 7.8 msec (sampling rate = 128 Hz). Furthermore, cross-correlations were computed between EEG signals and envelopes of randomly selected sentences which served as control condition (e.g., Horton et al., 2013). Any non-zero valued coefficients in these control functions were thus purely due to chance. Furthermore, the control condition we chose had the advantage of featuring the speaker's characteristics while not being temporally aligned to the respective EEG segment it was correlated with.

#### 2.7. Statistical analyses

##### 2.7.1. Determining the temporal windows for subsequent analyses

Testing how well the amplitude envelope was represented in the neural signal, we analysed the positive deflections in the cross-correlation functions (Abrams, Nicol, Zecker, & Kraus, 2008, 2009; Braiman et al., 2018), as values converging towards 1 indicate a positive linear relationship between the two signals. However, instead of extracting a single peak value, we compared each condition to the control condition for time lags ranging from 0 to 300 msec as these latency windows are known to include two prominent peaks in the



**Fig. 2 – Topographical distribution and time course of neural speech tracking. A: Topographical distribution and time course of the grand average cross-correlation in each listening condition from approximately 50 to 250 msec. B: Topographical distribution of the grand average cross-correlation at the peaks at approximately 100 and 200 msec. Selected electrode clusters are marked with “•”. Warm colours denote positive correlations, while cool colours denote negative correlations. C: Grand mean cross-correlation functions of the left and right temporo-parieto-occipital as well as the fronto-central electrode clusters. Significant time lags for the three conditions are marked as coloured bars at the bottom of each plot.**

cross-correlation function (e.g., Braiman et al., 2018; Horton et al., 2013; Zoefel & VanRullen, 2016), which were also apparent in our data (Fig. 2C). As opposed to Petersen et al. (2017) and in line with two previous studies (Braiman et al., 2018; Zoefel & VanRullen, 2016), we limited our analyses to these first two peaks, as the grand average neural tracking response in Babble appeared to “break down” in time lags of greater latency (Fig. 2C). To control for potential perceptual differences between the conditions, only correctly comprehended trials were extracted and compared with the control condition. As there were less correct trials in Babble, a random subsample of Quiet and Pink was extracted so that the same number of trials for each condition was included in the analysis.

For statistical analyses, electrodes in a left- (1-A10, 1-A11, 1-A12, 1-A13, 1-A14, 1-A15, 1-D24, 1-D25, 1-D26, 1-D30, 1-D31, 1-D32) and right-hemispheric temporo-parieto-occipital (1-A26, 1-A27, 1-A28, 1-B7, 1-B8, 1-B9, 1-B10, 1-B11, 1-B12, 1-

B14, 1-B15, 1-B16) as well as a fronto-central (1-B30, 1-B31, 1-B32, 1-C1, 1-C2, 1-C3, 1-C11, 1-C23, 1-C24, 1-D1, 1-D2, 1-D3, 1-D11, 1-D12, 1-D13)<sup>1</sup> cluster were selected (Fig. 2B). This selection was based on the topography and time course of the grand average cross-correlation signal of all conditions (Fig. 2A and B). The temporo-parieto-occipital clusters were divided into two separate electrode pools so that each cluster contained approximately the same number of electrodes and to investigate possible effects of lateralized processing. Note that the temporo-parieto-occipital clusters encompassed an early time window around 100 msec and the fronto-central cluster encompassed a later time window around 200 msec.

<sup>1</sup> BioSemi caps with  $\geq 128$  electrodes have an “ABC” layout that are according to the BioSemi designed equiradial system. The international 10–20 and the ABC layouts partly overlap but are not the same.

As expected, the control envelopes and the neural signal showed no systematic relationship with coefficients located around zero (Fig. 2C). We then ran paired-sample t-tests between each condition's and the control's grand average cross-correlation function for each time lag while significance levels were Bonferroni corrected, resulting in an alpha-level of .0002 (243 statistical tests: for each lag in each condition and each cluster). The average cross-correlation functions for each condition and cluster are depicted in Fig. 2C while significant time lags are highlighted. These time lags constituted the temporal window for all further analyses. Statistical tests were conducted in R (version 4.0.3; R Core Team, 2020).

### 2.7.2. Neural speech tracking and pure-tone hearing

The first aim of the present study was to investigate the effects of pure-tone hearing loss on neural speech tracking in listening conditions of varying adversity and in spatially different surface clusters and time windows respectively. Given the dependencies among data points (i.e., each subject completed several trials in each condition and data were collected from each cluster) it was deemed appropriate to fit a linear mixed-effects model (LMM). LMMs offer a powerful alternative to often used procedures such as (repeated measures) ANOVAs that rely on averaging techniques. Performed on trial-level data, LMMs take all data into account and are robust to unbalanced datasets (see Frömer, Maier, and Abdel Rahman (2018) for an overview of the advantages of (G) LMMs in EEG research). Note that in this model we analysed data from correctly comprehended trials only (i.e., the same trials that we used to determine the significant time windows). First, the model with the maximal random effects structure justified by the design was estimated (Barr, Levy, Scheepers, & Tily, 2013) using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). The model included fixed effects of PTA (continuous variable: in dB HL), condition (categorical variable with three levels: Quiet, Pink, Babble), cluster (categorical variable with three levels: Frontal, Left, Right), the three-way interaction between PTA, condition and cluster as well as the covariate age (continuous variable: in years). The continuous predictors PTA and age were z-standardized. As for the random effects, the model included by-subject and by-item random intercepts (i.e., crossed random effects; Baayen, Davidson, & Bates, 2008). Following the general rule of fitting a random slope for each focal within-unit predictor, the model contained by-subject random slopes for condition, cluster, their interaction as well as by-item random slopes for PTA, condition and cluster and their interaction. The random effects structure was further simplified by iteratively removing terms until a non-singular fit was achieved (Barr et al., 2013) to prevent the fitting of a model that was possibly degenerate or overparameterized (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017; Singmann & Kellen, 2019). Using the formula notation in R, the maximal model was defined as follows:

$$\text{crosscorr} \sim 1 + \text{PTAZ} * \text{condition} * \text{cluster} + \text{ageZ} + (1 + \text{condition} * \text{cluster} | \text{subject}) + (1 + \text{PTAZ} * \text{condition} * \text{cluster} | \text{item})$$

Because of the presence of higher-order effects (i.e., interactions), the default contrast coding scheme (i.e., treatment contrasts) was changed to an orthogonal sum-to-zero coding

system (i.e., effects or deviation coding) (Singmann & Kellen, 2019). The lower-level effects (i.e., main effects) were thus estimated at the level of the grand mean and interpreted accordingly. Statistical inference was drawn by using a step-down model building approach whereat likelihood ratio tests were performed by comparing an encompassing model with a reduced model that omitted the fixed effects of interest. The final model identified by the step-down approach was fitted using restricted maximum likelihood (REML). We used Satterthwaite's method for approximating degrees of freedom for the t-tests and corresponding *p*-values as implemented in the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017). The covariate age was kept in the final model despite not reaching statistical significance. Post hoc, pairwise comparisons between the condition and electrode cluster levels were computed using the *emmeans* package (Lenth, Singmann, Love, Buerkner, & Herve, 2018) while the initially set alpha level of .05 was adjusted using Tukey's test for multiple comparisons.

### 2.7.3. Neural speech tracking and behavioural performance

Another goal was the investigation of potential associations between neural speech tracking and behavioural performance in an intelligibility and a comprehension task. For this purpose, we compared neural tracking of speech between correct and incorrect trials by fitting two generalized linear mixed-effects models (GLMM) (i.e., one with intelligibility and one with comprehension as outcome measure) specified as binomial model with a logistic link function. The binary response measures were coded as 1 and 0 based on the participants' responses (1 = correct; 0 = incorrect). The two models contained fixed effects of PTA, cross-correlation and condition, the three-way interaction between these variables, cluster as control variable of non-primary interest and the covariate age. Continuous predictors were again z-standardized. As above, the random effects structure was determined by reducing the maximal model until a non-singular fit was reached. The maximal model took the following configuration:

$$\text{response} \sim 1 + \text{PTAZ} * \text{crosscorrZ} * \text{condition} + \text{cluster} + \text{ageZ} + (1 + \text{condition} | \text{subject}) + (1 + \text{PTAZ} * \text{crosscorrZ} * \text{condition} | \text{item})$$

To establish the significance of all main effects and interactions, likelihood ratio tests were again performed by comparing an encompassing model with a reduced model that omitted the fixed effects of interest. As above, orthogonal contrasts were used as factor coding and post hoc pairwise comparisons between all condition levels were computed.

## 3. Results

### 3.1. Significant time windows

We computed cross-correlations between EEG time series and the amplitude envelope of spoken sentences. As in previous work (Horton et al., 2013; Zoefel & VanRullen, 2016), the grand average cross-correlation functions showed a prominent peak at ~ 100 msec and a later one at around ~ 200 msec (Fig. 2C)

with an inversed polarity across the scalp (Fig. 2). It should be noted that envelope tracking responses in these time windows primarily stem from sources in auditory areas (e.g., Fiedler, Wöstmann, Herbst, & Obleser, 2019). These effects would have been maximized at frontal electrodes if mastoid electrodes were used as reference. However, as we did not record mastoids and as we used an average reference, the effects observed at the parieto-occipital and fronto-central electrodes reflect the same effects with an inversion in polarity (for a discussion on the influence of different reference schemes on brain responses see Luck, 2014).

Pairwise comparisons between the cross-correlation functions and the control condition revealed several significant time lags at a Bonferroni corrected alpha-level of .0002 (Fig. 2C). In the left temporo-parieto-occipital electrode cluster significant lags were found from 78 to 125 msec for Quiet, from 70 to 133 msec for Pink and from 102 to 156 msec for Babble. In the right temporo-parieto-occipital cluster, lags from 40 to 117 msec appeared to be significant for Quiet, whereat lags from 16 to 117 msec and from 47 to 133 msec were significant for Pink and Babble. In the fronto-central cluster, time lags from 141 to 242 msec were significant for Quiet whereat lags ranging from 148 to 258 msec and from 180 to 297 msec were significant for Pink respectively Babble.

### 3.2. Neural speech tracking as a function of pure-tone hearing and background noise

From these established time windows we extracted trial-level data by averaging cross-correlation values for each trial over all significant time lags within each condition and cluster. We then fitted an LMM to estimate the effects of pure-tone hearing on neural speech tracking. Furthermore, we examined differences between listening conditions of varying difficulty and between different electrode clusters. As for the fixed effects, none of the interactions between PTA, condition and cluster contributed to the explained variance in neural speech tracking ( $p > .05$ ) and were therefore excluded from the model. Furthermore, our analyses revealed a significant main effect of PTA ( $\Delta\chi^2(1) = 5.85, p = .016$ ), with a higher degree of pure-tone hearing loss being associated with enhanced neural speech tracking (Table 1; Fig. 3A). In addition, we found a significant main effect of condition ( $\Delta\chi^2(2) = 12.27, p = .002$ ) with a lower tracking response in Babble compared to the grand mean whereas this difference was not significant for Pink (Table 1). As we were interested in all contrasts possible, post hoc pairwise comparisons were computed, revealing

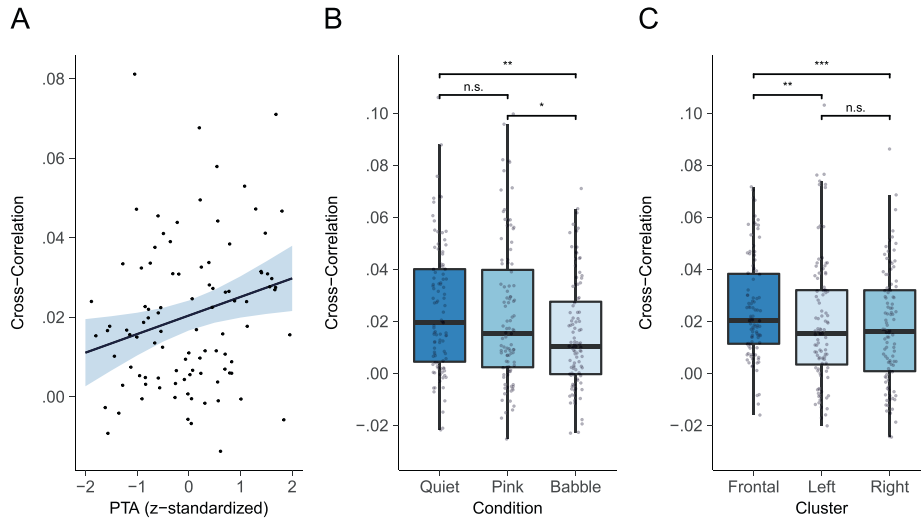
lower speech tracking in Babble than in Pink (Babble - Pink:  $\Delta\beta = -.008, SE = .003, t = -2.61, p = .028$ ) and in Babble compared to Quiet (Babble - Quiet:  $\Delta\beta = -.010, SE = .003, t = -3.29, p = .004$ ) whereas Pink and Quiet showed no significant difference (Pink - Quiet:  $\Delta\beta = -.002, SE = .003, t = -.62, p = .530$ ) (Fig. 3B). Additionally, we found significant differences between the electrode clusters under investigation ( $\Delta\chi^2(2) = 16.39, p < .001$ ). Pairwise comparisons between the three clusters revealed a significantly higher speech tracking response in the fronto-central compared to the left (Left - Frontal:  $\Delta\beta = -.005, SE = .002, t = -2.36, p = .020$ ) as well as the right temporo-parieto-occipital electrode cluster (Right - Frontal:  $\Delta\beta = -.007, SE = .002, t = -4.07, p < .001$ ) (Fig. 3C). However, there was no significant difference between the left and right temporo-parieto-occipital cluster (Left - Right:  $\Delta\beta = .002, SE = .002, t = 1.02, p = .310$ ).

Our results regarding the association between pure-tone hearing and neural speech tracking (Table 1; Fig. 3A) are in agreement with several other studies (Decruy et al., 2020; Fuglsang et al., 2020; Gillis et al., 2021; Mirkovic et al., 2019), but contradicting to the findings by Petersen et al. (2017). This discrepancy is most likely attributable to methodological differences: Petersen et al. (2017) computed the correlation between participants PTA and the neural tracking response in a limited time window, roughly corresponding to the window covered by our fronto-central cluster. We suspected that pure-tone hearing loss might exert its impact on speech-brain alignments at an earlier processing stage. Although the non-significant interaction between PTA and cluster in our model suggested otherwise, we tested this assumption directly by extracting data from the two temporo-parieto-occipital clusters (encompassing a time window around 100 msec) and the fronto-central cluster (encompassing a later time window around 200 msec) and reran two separate LMMs with the same fixed effects as described in Table 1 (except that the fixed effect term for cluster was omitted in the model comprising data from the fronto-central cluster). As for the random effects, we proceeded as described above (see section 2.7.2.) in order to find the model with the final random effects structure that did not produce a singular fit. Indeed, only the model that included data from the left and right temporo-parieto-occipital cluster, and thus the earlier time window, showed a significant and comparable effect of pure-tone hearing on neural speech tracking ( $\beta = .0054, t(109.09) = 2.62, p = .010$ ) (Table S1) while this effect for the later time window, indicated by the fronto-central electrodes, was not significant ( $\beta = .0026, t(100.41) = 1.39, p = .169$ ) (Table S2).

**Table 1 – Parameter estimates for the effects of PTA, condition and cluster on neural speech tracking. Model configuration:  $\text{crosscorr} \sim 1 + \text{PTAZ} + \text{condition} + \text{cluster} + \text{ageZ} + (1 + \text{condition} + \text{cluster} | \text{subject}) + (1 + \text{PTAZ} + \text{condition} | \text{item})$ .**

Predictors	Estimate	SE	df	t	p
Intercept	.0204	.0019	124.20	10.63	<.001
PTAZ	.0047	.0019	118.20	2.47	.015
condition (Babble)	-.0059	.0017	127.10	-3.54	<.001
condition (Pink)	.0019	.0018	117.90	1.05	.296
cluster (Left)	-.0009	.0014	97.49	-.64	.526
cluster (Right)	-.0033	.0012	89.38	-2.70	.008
ageZ	-.0014	.0018	96.74	-.80	.428





**Fig. 3 – Neural speech tracking as a function of hearing, condition and cluster. A: Predicted marginal effect of PTA on neural speech tracking. Shaded area depicts the 95% confidence interval. B: Condition differences in neural speech tracking. C: Cluster differences in neural speech tracking. To facilitate visual representation, trial-level data were averaged so that each dot represents a single participant. n.s. = not significant, \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .**

**3.3. Neural speech tracking and behavioural performance**

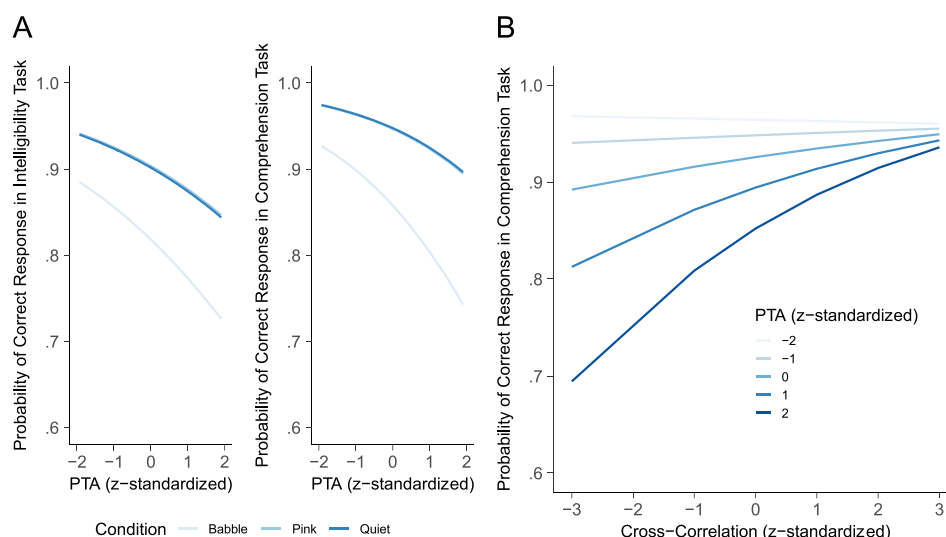
In a further step we assessed whether behavioural performance in an intelligibility and a comprehension task was related to the degree of neural speech tracking and whether there was a difference between the conditions. Intelligibility scores were significantly above chance (one-sample  $t$ -tests against chance (50%): Quiet,  $t(100) = 20.74, p < .001$ ; Pink,  $t(100) = 9.76, p < .001$ ; Babble,  $t(100) = 10.99, p < .001$ ) with a mean performance of 75.63% ( $\pm 1.24\%$  SEM) in Quiet, 68.80% ( $\pm 1.93\%$  SEM) in Pink and 67.61% ( $\pm 1.60\%$  SEM) in Babble. In order to estimate the difference in neural speech tracking between correct and incorrect trials, two GLMMs were fitted. As for the intelligibility task, likelihood ratio tests indicated no statistically significant effect of the three-way interaction between PTA, cross-correlation and condition ( $\Delta\chi^2(7) = 6.76, p = .454$ ), nor the interactions between PTA and cross-correlation ( $\Delta\chi^2(1) = .31, p = .579$ ), PTA and condition ( $\Delta\chi^2(2) = .89, p = .640$ ) or cross-correlation and condition ( $\Delta\chi^2(2) = 5.21, p = .074$ ). Regarding the main effects, the analyses revealed no significant relationship between our

measure of speech tracking and performance ( $\Delta\chi^2(1) = .11, p = .743$ ) while the main effects of PTA ( $\Delta\chi^2(1) = 7.21, p = .007$ ) and condition ( $\Delta\chi^2(2) = 8.43, p = .015$ ) were significant (Table 2; Fig. 4A). Post hoc pairwise comparisons revealed no significant difference between Quiet and Pink (Quiet - Pink:  $\Delta\text{OR} = -.02, \text{SE} = .40, p = .998$ ) as well as Pink and Babble (Pink - Babble:  $\Delta\text{OR} = .73, \text{SE} = .36, p = .103$ ). However, the odds of responding correctly appeared to be significantly higher in Quiet compared to Babble (Quiet - Babble:  $\Delta\text{OR} = .71, \text{SE} = .27, p = .025$ ) (Fig. 4A). Regarding the effect of PTA, higher hearing thresholds were associated with worse task performance where a unit change in PTA (i.e., +1 standard deviation) was associated with a decrease of 24% in the odds of responding correctly (Table 2; Fig. 4A).

Subsequently, we investigated the association between neural speech tracking and speech comprehension. First of all, comprehension scores were significantly above chance (one-sample  $t$ -tests against chance (25%): Quiet,  $t(100) = 27.39, p < .001$ ; Pink,  $t(100) = 29.20, p < .001$ ; Babble,  $t(100) = 25.97, p < .001$ ) with a mean performance of 82.18% ( $\pm 2.09\%$  SEM) in Quiet, 82.67% ( $\pm 1.98\%$  SEM) in Pink and 75.08% ( $\pm 1.93\%$  SEM) in Babble. As for the GLMM, likelihood ratio tests indicated no

**Table 2 – Parameter estimates for the effects of PTA, cross-correlation and condition on performance in the intelligibility task. Model configuration: response ~ 1 + PTAZ + crosscorrZ + condition + cluster + ageZ + (1 + condition | subject) + (1 + PTAZ + crosscorrZ + condition | item).**

Predictors	Estimate	SE	Odds Ratio	$p$
Intercept	1.99	.14	7.29	<.001
PTAZ	-.28	.10	.76	.007
crosscorrZ	-.01	.03	.99	.736
condition (Babble)	-.48	.17	.62	.004
condition (Pink)	.25	.24	1.29	.288
cluster (Left)	.001	.03	1.00	.974
cluster (Right)	-.003	.03	1.00	.937
ageZ	.002	.08	1.00	.981



**Fig. 4 – Speech tracking and behaviour. A: Predicted probability of a correct response in the intelligibility (left) and the comprehension (right) task as a function of PTA and condition. Note that the regression lines of Quiet and Pink overlap. B: Interaction between PTA and cross-correlation on predicted probability of a correct response in the comprehension task. As the interaction contains two continuous predictors, each line represents the predicted probability for participants with 5 different PTA levels (–2, –1, 0, 1 and 2 Standard deviations where 0 depicts the mean PTA).**

statistically significant effect of the three-way interaction between PTA, cross-correlation and condition ( $\Delta\chi^2(7) = 6.71$ ,  $p = .459$ ), nor the interactions between PTA and condition ( $\Delta\chi^2(2) = .02$ ,  $p = .991$ ), or cross-correlation and condition ( $\Delta\chi^2(2) = .36$ ,  $p = .837$ ). The interaction between PTA and cross-correlation on the other hand was significant ( $\Delta\chi^2(1) = 4.69$ ,  $p = .030$ ) (Table 3). As can be seen in Fig. 4B, the higher the neural tracking response, the more likely it was to comprehend a sentence correctly but only in participants with bad hearing. In better hearing participants this relationship diminished. In other words, speech tracking seemed to benefit comprehension, especially in participants with very bad pure-tone hearing. This is also indicated by the odds ratio in Table 3 which shows that a unit change in cross-correlation (i.e., +1 standard deviation) was associated with an 8% increase in the odds of answering correctly with PTAZ held at 0 (i.e., the mean PTA). In a better hearing participant (e.g., with a PTA one

standard deviation below the mean) a unit change in cross-correlation was associated with a 3% increase in the odds of answering correctly ( $1.08^{-1} * 1.12$ ) whereas in a worse hearing participant (e.g., with a PTA one standard deviation above the mean) the odds increased by 21% ( $1.08^1 * 1.12$ ) with a unit change in cross-correlation.

As for the main effects, there appeared to be significant effects of PTA ( $\Delta\chi^2(1) = 6.97$ ,  $p = .008$ ), cross-correlation ( $\Delta\chi^2(1) = 5.98$ ,  $p = .015$ ) and condition ( $\Delta\chi^2(2) = 23.62$ ,  $p < .001$ ). As can be seen in Table 3, a one unit increase in cross-correlation was associated with an increase of 12% in the odds of responding correctly, whereas a unit increase in PTA led to a decrease in the odds of 49%. Post hoc pairwise comparisons between the three conditions revealed lower odds to respond correctly in Babble compared to Pink (Pink - Babble:  $\Delta\text{OR} = 2.93$ ,  $\text{SE} = .72$ ,  $p < .001$ ) and in Babble compared to Quiet (Quiet - Babble:  $\Delta\text{OR} = 1.83$ ,  $\text{SE} = .59$ ,  $p = .006$ ), whereas Quiet and Pink showed no significant difference (Quiet - Pink:  $\Delta\text{OR} = -1.10$ ,  $\text{SE} = .83$ ,  $p = .377$ ) (Fig. 4B).

We followed up our analyses by further exploring the main effect of cross-correlation and the interaction between PTA and cross-correlation. We took advantage of the temporal information inherent to the cross-correlation function by running GLMMs for the two temporo-parieto-occipital and the fronto-central clusters separately. By doing so, we aimed to investigate whether the observed effects were restricted to the earlier or the later time window. The two models had the same configuration with respect to the fixed effects as depicted in Table 3 (except that the fixed effect term for cluster was omitted in the model comprising data from the fronto-central cluster). As for the random effects, we again proceeded as described above (see section 2.7.2.). The results suggest that the effect of cross-correlation was mainly present in the later time window with a one unit increase in cross-correlation (i.e., 1 standard deviation) being related to an increase of 15% in the

**Table 3 – Parameter estimates for the effects of PTA, cross-correlation and condition on performance in the comprehension task. Interactions are indicated by the symbol “:”. Model configuration: response ~ 1 + PTAZ \* crosscorrZ + condition + cluster + ageZ + (1 + condition | subject) + (1 + PTAZ + crosscorrZ + condition | item).**

Predictors	Estimate	SE	Odds Ratio	p
Intercept	4.44	.36	84.69	<.001
PTAZ	-.68	.26	.51	.009
crosscorrZ	.11	.05	1.12	.021
condition (Babble)	-1.59	.34	.20	<.001
condition (Pink)	1.34	.48	3.83	.005
cluster (Left)	.005	.04	1.00	.917
cluster (Right)	.001	.04	1.00	.977
ageZ	.44	.25	1.55	.074
PTAZ:crosscorrZ	.08	.04	1.08	.031

odds of responding correctly (Table S4). Although not statistically significant the interaction effect appears to be greater in the fronto-central compared to the temporo-parieto-occipital clusters (OR: 1.09 vs 1.05) (Tables S3 and S4) suggesting that this interaction was indeed mainly present in the later time window.

In a further step we explored whether the effects of the between-subject predictors PTA and cross-correlation on performance in the comprehension task were subject to a potential confound (i.e., due to the Simpson's paradox). Specifically, we wanted to test whether the effects we found were present when omitting the Babble condition. Given the reduced neural speech tracking in Babble it was unclear whether including Babble in the statistical model might have introduced a potential confound. We therefore fitted a GLMM using the same configuration as described in Table 3 while including data from Quiet and Pink only. The results suggest that the effects were comparable to the model that included all levels of condition for PTA (OR: .52 vs .51) and cross-correlation (OR: 1.10 vs 1.12). We therefore concluded that the inclusion of Babble in the statistical model did not confound the effects we observed.

---

#### 4. Discussion

In the current study, we investigated the relationship between uncorrected pure-tone hearing loss and neural tracking of the slow varying temporal envelope of speech in a large sample of older adults. Furthermore, we were interested in the effects of masking on the envelope tracking response. Lastly, we examined the role of neural speech tracking for behavioural performance in an intelligibility and a comprehension task. For this purpose, a large sample of older adults with varying degrees of hearing loss listened to spoken sentences in quiet and in the presence of interfering background noise while their neurophysiological activity was recorded using scalp EEG. To get a measure of how well the acoustic signal was represented in the neural data, we computed the cross-correlation between the two time series.

Our results revealed a linear increase in neural speech tracking (i.e., greater cross-correlation) as a function of pure-tone hearing loss, corroborating several other studies (Decruy et al., 2020; Fuglsang et al., 2020; Gillis et al., 2021; Mirkovic et al., 2019). In other words, the more severe the pure-tone hearing loss, the better the phase-alignment between acoustic envelope modulations and slow neural activity. This finding, however, does not fully correspond to the results from the studies by Presacco et al. (2019) and Petersen et al. (2017) where no such relationship had been found. Presacco et al. (2019) compared the envelope tracking response in two older groups of normal hearing and hearing-impaired adults and a young normal hearing control group. These authors found that envelope tracking did not differ between the two older groups. The deviance from the findings of Presacco et al. (2019) might be attributable to alterations in how the brain responds to acoustic stimulation as a function of hearing loss. In a recent review, Herrmann and Butler (2021) pointed out that current measures of hyperactivity in the human auditory cortex (i.e., responsivity to acoustic

stimulation measured as sound-evoked responses) are likely non-linear. To a certain level of hearing acuity, sound-evoked responses are enhanced in hearing-impaired compared to normal hearing adults. However, in severe hearing loss where cochlear damage is profound and signal transduction is seriously restricted, hyperresponsivity can no longer be measured. The authors emphasize that the absence of measurable hyperresponsivity does however not imply the absence of hyperactivity which poses one possible reason why some authors find hyperactivity as a function of hearing loss and others do not. Although not explicitly reported in their publication and therefore derived from their plotted audiogram, the hearing-impaired group in the study by Presacco et al. (2019) had a PTA of around 49 dB HL. Using the same criterion as Presacco and colleagues (PTA  $\geq$ 26 dB HL from .5 to 4 kHz), our subsample of hearing-impaired adults showed much better hearing levels (PTA = 33.7 dB HL). To underline the points made by Herrmann and Butler (2021), it might be productive to consider the non-linearity between hyperresponsivity and hearing acuity in future studies that investigate sound-evoked responses in clinical populations with impaired hearing, especially when hearing loss is severe.

The discrepancy between our findings and the findings of Petersen et al. (2017) on the other hand might be attributable to methodological differences. The authors computed the correlation between participants' PTA and the cross-correlation in a limited time window, roughly corresponding to the window covered by our fronto-central cluster. We hypothesized that pure-tone hearing loss might exert its impact on speech-brain alignment at an earlier processing stage. To test this assumption, we extracted data from two temporo-parieto-occipital clusters (encompassing an earlier time window around 100 msec) and a fronto-central cluster (encompassing a later time window around 200 msec) and reran two separate LMMs. Indeed, only the model that included data from the left and right temporo-parieto-occipital clusters showed a significant association between pure-tone hearing and neural speech tracking, indicating that the enhanced responses were restricted to the earlier time window.

However, in light of the speech perception hypothesis proposed by Giraud and Poeppel (2012) these findings seem counterintuitive. As these authors propose, slow oscillatory activity phase aligns to temporal regularities in the speech signal, segmenting the speech stream into meaningful chunks for subsequent processing. Considering that this initial chunking process lies at the core for spoken language comprehension, one might expect this process to be compromised or at least not enhanced in hearing-impaired adults. However, this reasoning is not in harmony with our observation. One possible conclusion would be that hearing-impaired adults more strongly rely on slow acoustic features in the speech signal transposed by the temporal envelope. Typically, these slow acoustic features carry prosodic or rhythmic information which has been demonstrated to be essential for older adults (Giroud, Keller, Hirsiger, Dellwo, & Meyer, 2019; Meyer, Keller, & Giroud, 2018). The commitment to slow acoustic modulations to some degree compensates for the reduced audibility that follows a pure-tone hearing loss. This assumption is strengthened by our finding, which showed that the association between speech tracking

and comprehension became stronger as hearing got worse. We therefore argue that, in light of the aforementioned speech perception framework, this initial stimulus encoding, as reflected by neural speech tracking, is in fact more precise in the hearing-impaired. We, however, also argue that it only accounts partly for the observed response enhancements. By using natural speech and determining its congruence with the neural signal it is most likely that, on top of tracking that reflects higher-level processing, we captured a series of evoked responses elicited by acoustic regularities in the speech signal (Obleser & Kayser, 2019; Zoefel, ten Oever, & Sack, 2018; for a study supporting the contribution of both processes to the measured signal see; Doelling, Assaneo, Bevilacqua, Pesaran, & Poeppel, 2019). As cochlear damage has been linked to hyperactivity along the auditory pathway (e.g., Herrmann & Butler, 2021), we assume that parts of the observed speech-brain synchronization that subserve a functional role for speech comprehension might be overshadowed by passively elicited responses that are per se higher in hearing-impaired listeners. Notwithstanding, it would be interesting to investigate whether the alignment between the regularity that is inherent to the high-level features of speech and low-frequency activity of the cortex is affected by age and pure-tone hearing loss. This has already been done in young normal hearing adults (Zoefel & VanRullen, 2015, 2016). The authors constructed intelligible speech stimuli that lacked systematic fluctuations in the amplitude envelope and spectral content and thus low-level features, while keeping the regularities in high-level acoustic and linguistic features (i.e., “features that do not passively entrain the lowest levels of auditory processing”; see Zoefel & VanRullen, 2015) intact. They showed that the neural tracking response was elicited despite the lack of low-level acoustic regularities (Zoefel & VanRullen, 2016). Interestingly, only a later peak (denoted as “high-level component”; corresponding to the time window encompassed by our fronto-central cluster) was induced by their constructed speech stimuli whereas an earlier one (denoted as “low-level component”; corresponding to the time-window encompassed by our two temporo-parieto-occipital clusters) was restricted to unprocessed speech. They speculated that the earlier peak was evoked by low-level acoustic features only apparent in unprocessed speech—a conjecture that provides further support for our argumentation outlined above. As the association between pure-tone hearing and neural speech tracking seemed to be mainly driven by processes in the earlier time window, it is likely that at least parts of the neural tracking enhancements in hearing-impaired listeners were due to brain responses to low-level acoustic features and thus not reflecting processes subserving speech comprehension.

The temporal distinction between “low- and high-level processing” also becomes relevant in connection with our behavioural analyses. We assessed whether neural speech tracking was related to behavioural performance in an intelligibility and a comprehension task in different conditions and found speech tracking to be positively associated with how well sentences were comprehended. This is in line with previous studies that showed the same relationship in normal hearing and hearing-impaired adults (Decruy et al., 2020; Gillis et al., 2021; Kurthen et al., 2021). Surprisingly, this association

was absent for the intelligibility task. We assume that the use of short sound snippets might not be suitable to assess intelligibility as it reflects pure pattern matching where an access to any lexical representation is absent. If we assume a functional role of speech tracking for spoken language comprehension, this very process might not be necessary for a task where sublexical processing is probed.

The positive association between neural speech tracking and comprehension further suggests that perceptual problems experienced by older people in difficult listening conditions are entailed by deteriorated neural processes. These very processes could therefore be used as an objective marker for speech comprehension (Decruy et al., 2019, 2020; Lesenfans, Vanthornhout, Verschuere, Decruy, & Francart, 2019; Vanthornhout, Decruy, Wouters, Simon, & Francart, 2018) or as a promising target mechanism for clinical interventions (e.g., as a target for training studies such as neurofeedback). As discussed above, our study additionally showed that the relationship between speech tracking and comprehension was modulated by participant's pure-tone hearing. Speech tracking seemed to be important for comprehension, especially in participants with impaired hearing, suggesting a compensatory mechanism. Interestingly, this effect seemed to be mainly driven by processes in a latency window at around 200 msec encompassed by the fronto-central electrodes which is in line with the study by Kurthen et al. (2021). To resume the discussion from above, these findings further underline the importance of considering the temporal dimension to distinguish passively elicited responses and speech tracking that plays an important role in the language processing hierarchy. Applying stimuli as in the studies by Zoefel and VanRullen (2015, 2016) in aging and hearing research could disentangle the contributions of low- and high-level processes to the observed speech-brain synchronization and thus give us an idea of the processing stage at which speech perception difficulties arise.

---

## 5. Conclusion

This study investigated the relationship between uncorrected pure-tone hearing loss and neural processing of continuous speech under different masking conditions. There was a linear increase in neural speech tracking as a function of pure-tone hearing loss. Furthermore, we found that masking had detrimental effects on behavioural task performance, which was also reflected in deteriorated neural tracking responses. Lastly, neural speech tracking was predictive for behavioural performance in a comprehension task. Interestingly, this relationship was modulated by the degree of pure-tone hearing loss, suggesting a compensatory mechanism whereby hearing-impaired adults rely more on the slow varying temporal envelope of speech.

---

## Data statement

The complete neural and behavioural data as well as the auditory stimuli used in this study are publicly available in the study's Open Science Framework repository ([https://osf.](https://osf)

io/s3kzq/). We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. No part of the study procedures and analyses was pre-registered prior to the research being conducted.

---

### Credit author statement

**Raffael Schmitt:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - Original Draft, Visualization; **Martin Meyer:** Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition; **Nathalie Giroud:** Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Funding acquisition.

---

### Open practices

The study in this article earned Open Data and Open Materials badges for transparent practices. Materials and data for the study are available at <https://osf.io/s3kzq/>.

---

### Funding

This work was supported by grants from the Swiss National Science Foundation (grant no. 105319\_169964 to MM and grant no. PR00P1\_185715 to NG), the “Forschungskredit” of the University of Zurich (grant no. FK-19-072 to NG) and the “Zürcher Stiftung für das Hören (ZSFH)”.

---

### Declaration of competing interest

The authors have no conflict of interest to declare.

---

### Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2022.02.017>.

---

### REFERENCES

- Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience*, 28(15), 3958–3965. <https://doi.org/10.1523/JNEUROSCI.0187-08.2008>
- Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience*, 29(24), 7686–7693. <https://doi.org/10.1523/JNEUROSCI.5242-08.2009>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Biesmans, W., Das, N., Francart, T., & Bertrand, A. (2017). Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(5), 402–412. <https://doi.org/10.1109/TNSRE.2016.2571900>
- Boersma, P., & Weenink, D. (2021). Praat: Doing phonetics by computer [Computer program]. (6.1.40).
- Braiman, C., Fridman, E. A., Conte, M. M., Voss, H. U., Reichenbach, C. S., Reichenbach, T., et al. (2018). Cortical response to the natural speech envelope correlates with neuroimaging evidence of cognition in severe brain injury. *Current Biology*, 28(23), 3833–3839. <https://doi.org/10.1016/j.cub.2018.10.057>
- Ciorba, A., Bianchini, C., Pelucchi, S., & Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical Interventions in Aging*, 7, 159–163. <https://doi.org/10.2147/CIA.S26059>
- Decruy, L., Vanthornhout, J., & 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. <https://doi.org/10.1152/jn.00687.2018>
- Decruy, L., Vanthornhout, J., & Francart, T. (2020). Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hearing Research*, 393, 107961. <https://doi.org/10.1016/j.heares.2020.107961>
- Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, 8, 311. <https://doi.org/10.3389/fnhum.2014.00311>
- Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, 85, 761–768. <https://doi.org/10.1016/j.neuroimage.2013.06.035>
- Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences of the United States of America*, 116(20), 10113–10121. <https://doi.org/10.1073/pnas.1816414116>
- Fiedler, L., Wöstmann, M., Herbst, S. K., & Obleser, J. (2019). Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *Neuroimage*, 186, 33–42. <https://doi.org/10.1016/j.neuroimage.2018.10.057>
- Frömer, R., Maier, M., & Abdel Rahman, R. (2018). Group-level EEG-processing pipeline for flexible single trial-based analyses including linear mixed models. *Frontiers in Neuroscience*, 12, 48. <https://doi.org/10.3389/fnins.2018.00048>
- Fuglsang, S. A., Märcher-Rørsted, J., Dau, T., & Hjørtkjær, J. (2020). Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. *Journal of Neuroscience*, 40(12), 2562–2572. <https://doi.org/10.1523/JNEUROSCI.1936-19.2020>
- Gates, G. A., & Mills, J. H. (2005). Presbycusis. *Lancet*, 366(9491), 1111–1120. [https://doi.org/10.1016/S0140-6736\(05\)67423-5](https://doi.org/10.1016/S0140-6736(05)67423-5)
- Ghitza, O. (2013). The theta-syllable: A unit of speech information defined by cortical function. *Frontiers in Psychology*, 4, 138. <https://doi.org/10.3389/fpsyg.2013.00138>

- Gillis, M., Decruy, L., Vanthornhout, J., & Francart, T. (2021). Hearing loss is associated with delayed neural responses to continuous speech. *BioRxiv*.
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517. <https://doi.org/10.1038/nn.3063>
- Giroud, N., Hirsiger, S., Muri, R., Kegel, A., Dillier, N., & Meyer, M. (2018). Neuroanatomical and resting state EEG power correlates of central hearing loss in older adults. *Brain Structure & Function*, 223, 145–163. <https://doi.org/10.1007/s00429-017-1477-0>
- Giroud, N., Keller, M., Hirsiger, S., Dellwo, V., & Meyer, M. (2019). Bridging the brain structure – brain function gap in prosodic speech processing in older adults. *Neurobiology of Aging*, 80, 116–126. <https://doi.org/10.1016/j.neurobiolaging.2019.04.017>
- Giroud, N., Lemke, U., Reich, P., Matthes, K. L., & Meyer, M. (2017). The impact of hearing aids and age-related hearing loss on auditory plasticity across three months – an electrical neuroimaging study. *Hearing Research*, 353, 162–175. <https://doi.org/10.1016/j.heares.2017.06.012>
- Herrmann, B., & Butler, B. E. (2021). Hearing loss and brain plasticity: The hyperactivity phenomenon. *Brain Structure & Function*, 226(7), 2019–2039. <https://doi.org/10.1007/s00429-021-02313-9>
- Homans, N. C., Metselaar, R. M., Dingemans, J. G., van der Schroeff, M. P., Brocaar, M. P., Wieringa, M. H., et al. (2017). Prevalence of age-related hearing loss, including sex differences, in older adults in a large cohort study. *The Laryngoscope*, 127(3), 725–730. <https://doi.org/10.1002/lary.26150>
- Horton, C., D'Zmura, M., & Srinivasan, R. (2013). Suppression of competing speech through entrainment of cortical oscillations. *Journal of Neurophysiology*, 109(12), 3082–3093. <https://doi.org/10.1152/jn.01026.2012>
- Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., et al. (2012). Central presbycusis: A review and evaluation of the evidence. *Journal of the American Academy of Audiology*, 23(8), 635–666. <https://doi.org/10.3766/jaaa.23.8.5>
- de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41(2), 385–390. <https://doi.org/10.3758/BRM.41.2.385>
- Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2), 163–178. <https://doi.org/10.1017/S0048577200980259>
- Kurthen, I., Galbier, J., Jagoda, L., Neuschwander, P., Giroud, N., & Meyer, M. (2021). Selective attention modulates neural envelope tracking of informationally masked speech in healthy older adults. *Human Brain Mapping*, 42(10), 3042–3057. <https://doi.org/10.1002/hbm.25415>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13). <https://doi.org/10.18637/jss.v082.i13>
- Lawrence, B. J., Jayakody, D. M. P., Bennett, R. J., Eikelboom, R. H., Gasson, N., & Friedland, P. L. (2020). Hearing loss and depression in older adults: A systematic review and meta-analysis. *The Gerontologist*, 60(3), e137–e154. <https://doi.org/10.1093/geront/gnz009>
- Lecluyse, W., & Meddis, R. (2009). A simple single-interval adaptive procedure for estimating thresholds in normal and impaired listeners. *The Journal of the Acoustical Society of America*, 126(5), 2570–2579. <https://doi.org/10.1121/1.3238248>
- Lecluyse, W., Tan, C. M., McFerran, D., & Meddis, R. (2013). Acquisition of auditory profiles for good and impaired hearing. *International Journal of Audiology*, 52(9), 596–605. <https://doi.org/10.3109/14992027.2013.796530>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. In *R package version 1.1*.
- Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., & Francart, T. (2019). Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. *Hearing Research*, 380, 1–9. <https://doi.org/10.1016/j.heares.2019.05.006>
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Internal Medicine*, 173(4), 293–299. <https://doi.org/10.1001/jamainternmed.2013.1868>
- Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S., & Lawlor, B. A. (2018). Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis. *JAMA Otolaryngology - Head and Neck Surgery*, 144(2), 115–126. <https://doi.org/10.1001/jamaoto.2017.2513>
- Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). The MIT Press.
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001–1010. <https://doi.org/10.1016/j.neuron.2007.06.004>
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- Meyer, M., Keller, M., & Giroud, N. (2018). Suprasegmental speech prosody and the human brain. In S. Fruhholz, & P. Belin (Eds.), *The oxford handbook of voice perception* (pp. 143–165). Oxford University Press.
- Mirkovic, B., Debener, S., Schmidt, J., Jaeger, M., & 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. <https://doi.org/10.1016/j.heares.2019.04.005>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926. <https://doi.org/10.1016/j.tics.2019.08.004>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011. <https://doi.org/10.1155/2011/156869>
- Peelle, J. E., & Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends in Neurosciences*, 39(7), 486–497. <https://doi.org/10.1016/j.tins.2016.05.001>
- Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of Neurophysiology*, 117(1), 18–27. <https://doi.org/10.1152/jn.00527.2016>
- Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 21, 322–334. <https://doi.org/10.1038/s41583-020-0304-4>
- Presacco, A., Simon, J. Z., & Anderson, S. (2016). Evidence of degraded representation of speech in noise in the aging midbrain and cortex. *Journal of Neurophysiology*, 116(5), 2346–2355. <https://doi.org/10.1152/jn.00372.2016>

- Presacco, A., Simon, J. Z., & Anderson, S. (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *Plos One*, 14(3), Article e0213899. <https://doi.org/10.1371/journal.pone.0213899>
- R Core Team. (2020). R: A language and environment for statistical computing. In R: *A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2018). Neural entrainment to speech modulates speech intelligibility. *Current Biology*, 28(2), 161–169. <https://doi.org/10.1016/j.cub.2017.11.033>
- Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. Spieler, & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). Routledge. <https://doi.org/10.4324/9780429318405-2>.
- Søndergaard, P. L., & Majdak, P. (2013). The auditory modeling toolbox. In J. Blauert (Ed.), *The technology of binaural listening* (pp. 33–56). Springer. [https://doi.org/10.1007/978-3-642-37762-4\\_2](https://doi.org/10.1007/978-3-642-37762-4_2).
- Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends in Amplification*, 6(4), 131–165. <https://doi.org/10.1177/108471380200600402>
- Teng, X., Tian, X., Rowland, J., & Poeppel, D. (2017). Concurrent temporal channels for auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. *Plos Biology*, 15(11), Article e2000812. <https://doi.org/10.1371/journal.pbio.2000812>
- Vanthonhout, J., Decruy, L., Wouters, J., Simon, J. Z., & Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *JARO - Journal of the Association for Research in Otolaryngology*, 19, 181–191. <https://doi.org/10.1007/s10162-018-0654-z>
- Wilsch, A., Neuling, T., Obleser, J., & Herrmann, C. S. (2018). Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *Neuroimage*, 172, 766–774. <https://doi.org/10.1016/j.neuroimage.2018.01.038>
- World Health Organization. (2017). *Global costs of unaddressed hearing loss & cost-effectiveness of interventions: A WHO report, 2017* <https://apps.who.int/iris/bitstream/handle/10665/254659/9789241512046-eng.pdf>.
- Zoefel, B., Archer-Boyd, A., & Davis, M. H. (2018). Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. *Current Biology*, 28(3), 401–408. <https://doi.org/10.1016/j.cub.2017.11.071>
- Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience*, 12, 95. <https://doi.org/10.3389/fnins.2018.00095>
- Zoefel, B., & VanRullen, R. (2015). Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *Journal of Neuroscience*, 35(5), 1954–1964. <https://doi.org/10.1523/JNEUROSCI.3484-14.2015>
- Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *Neuroimage*, 124, 16–23. <https://doi.org/10.1016/j.neuroimage.2015.08.054>