

Hearing-Aid Directionality Improves Neural Speech Tracking in Older Hearing-Impaired Listeners

Trends in Hearing
Volume 26: 1–14
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165221099894
journals.sagepub.com/home/tia



Eline Borch Petersen 

Abstract

In recent years, a growing body of literature has explored the effect of hearing impairment on the neural processing of speech, particularly related to the neural tracking of speech envelopes. However, only limited work has focused on the potential usage of the method for evaluating the effect of hearing aids designed to amplify and process the auditory input provided to hearing-impaired listeners. The current study investigates how directional sound processing in hearing-aids, denoted directionality, affects the neural tracking and encoding of speech in EEG recorded from 11 older hearing-impaired listeners. Behaviorally, the task performance improved when directionality was applied, while subjective ratings of listening effort were not affected. The reconstruction of the to-be-attended speech envelopes improved significantly when applying directionality, as well as when removing the background noise altogether. When inspecting the modelled response of the neural encoding of speech, a faster transition was observed between the early bottom-up response and the later top-down attentional-driven responses when directionality was applied. In summary, hearing-aid directionality affects both the neural speech tracking and neural encoding of to-be-attended speech. This result shows that hearing-aid signal processing impacts the neural processing of sounds and that neural speech tracking is indicative of the benefits associated with applying hearing-aid processing algorithms.

Keywords

hearing aids, hearing loss, EEG, neural tracking of speech, listening effort

Received 22 June 2021; Revised received 18 March 2022; accepted 21 April 2022

Introduction

The ability to selectively attend to speech in the presence of background noise is crucial to successful communication, allowing listeners to separate individual speakers from the background noise and attend to them. About a decade ago, it was discovered that the neural activity of the listener reflects the spectro-temporal content of the speech being presented, and that the speech signal being attended to was better represented than speech being ignored (Ding & Simon, 2012; Howard & Poeppel, 2010; Mesgarani & Chang, 2012). For researchers focusing on the impaired hearing system, these results offered an interesting opportunity to investigate how the loss of hearing, and consequent reduced ability to selectively attend to individual speakers (Shinn-Cunningham & Best, 2008), affects neural speech tracking.

Unfortunately, the growing body of research within this field does not provide a unified answer to how impaired hearing affects neural tracking of the speech envelope. Initial studies of older adult listeners found that poorer hearing was associated with more faithful neural tracking

of to-be-ignored speech, i.e. a reduced ability to ignore irrelevant sounds (Dai, Best, & Shinn-Cunningham, 2018; Petersen et al., 2017). However, more recent studies have reported enhanced neural speech tracking of to-be-attended speech for hearing-impaired listeners compared to age-matched normal-hearing controls (Fuglsang et al., 2020; Millman et al., 2017). Although Goossens and colleagues identified similar enhanced neural tracking in hearing-impaired listeners compared with younger and middle-aged normal-hearing listeners, no difference was found between older listeners with normal and impaired hearing (Goossens et al., 2019). Similarly, Presacco and colleagues found no differences in neural speech tracking of older normal-hearing and hearing-impaired listeners (Presacco et al., 2019).

WS Audiology A/S, Lyngø, Denmark

Corresponding author:

Eline Borch Petersen, WS Audiology A/S, Lyngø, Denmark.
Email: eline.petersen@wsa.com



These contradicting effects of hearing impairment on neural speech tracking could arise from the complex interplay between the degree of hearing loss being affected by increased age, age affecting general neural processing and cognitive abilities (Grady, 2012; Grady et al., 2006), and age, cognitive abilities, and hearing loss in turn affecting speech processing (Dubno et al., 1984; Lunner et al., 2009). Further, the outcomes of the studies investigating effects of hearing loss on neural speech tracking might be affected by participant demographics, as well as the experimental design choices. These choices include questions such as: (1) Was the loss of hearing compensated for by hearing aids (Petersen et al., 2017), by digitally compensating the presented sound files (Millman et al., 2017), by increasing the overall loudness (Dai et al., 2018; Fuglsang et al., 2020; Goossens et al., 2019), or not at all (Presacco et al., 2019)? (2) If presenting background noise, was it individually adjusted to ensure equal speech intelligibility between listeners (Fuglsang et al., 2020; Petersen et al., 2017) or not (Millman et al., 2017; Presacco et al., 2019)? (3) Were differences observed with the degree of hearing impairment (correlation analysis; Petersen et al., 2017), or between age-matched groups with a narrow (Fuglsang et al., 2020; Millman et al., 2017; Presacco et al., 2019) or large age-span (Dai et al., 2018; Goossens et al., 2019)?

Although this body of research currently does not provide a clear indication of what happens to the brain's ability to track speech, it does reveal that the ability is not lost because of impaired hearing. Hence, it is possible to investigate what can be done for people suffering from impaired hearing to aid them in the listening situations requiring selective attention to a speaker in the presence of noise.

Modern hearing aids compensate for hearing impairment by presenting amplified and processed sounds to the wearer. In situations with multiple speakers, the hearing aids cannot currently identify which speaker is attended to by the wearer but can instead estimate the spectral content and spatial position of potential noise sources in order to attenuate them. Traditionally, the benefits of hearing-aid processing are evaluated using standardized speech-in-noise tests, requiring the participants to repeat sentences or words presented in noise. The outcome of these tests can be a performance score of the number of correctly repeated word/sentences, or a signal-to-noise level (SNR) at which a certain level of performance is obtained (Taylor, 2003). In either case, the outcome is a single value, not describing the process of listening in much detail. To combat this narrow view of evaluating speech understanding purely on intelligibility, an increasing number of studies focus on measuring the listening effort perceived by the listener using different biological markers. One of the most popular measures is that of pupillometry, which has shown reduced in listening effort, observed by a reduced pupil dilation, when hearing-aid noise reduction is applied in listening situations with close to perfect speech intelligibility (Wendt et al., 2017). More insights can be

gained if considering changes in the neural responses related to speech processing. The neural effects of hearing-aid signal processing have been investigated in a few studies: In a series of papers, Bernarding and colleagues found that applying hearing-aid noise reduction and directional sound processing, reduced the listening effort quantified by reduced entropy in the phase of the alpha-band activity (Bernarding et al., 2012; Bernarding et al., 2014, 2017). Alpha power has been found to increase with higher auditory processing load, e.g. by degrading the auditory signal, increasing the memory load, or altering the task complexity (Strauß et al., 2014). Alpha power has been observed to decrease when listeners were provided with narrow directional sound processing, compared to a more broad directivity pattern (Winneke et al., 2020). However, in an experiment testing the effect of hearing-aid noise reduction (directional processing in combination with frequency-specific noise reduction), no effect of its application was observed on the alpha power during listening (Fiedler et al., 2021). While on the same data, the hearing-aid noise reduction have been found to improve the neural speech tracking of both to-be-attended and to-be-ignored speech, while reducing the tracking of the background noise (Alickovic et al., 2020). In a study investigating the effect of directional sound processing, a cardioid directional pattern with a generic head-related impulse response was digitally applied to stimuli that were linearly amplified and presented to normal-hearing and hearing-impaired participants using insert-earphones (Mirkovic et al., 2019). Mirkovic and colleagues found that applying directional processing caused faster and improved neural encoding of the to-be-attended speech as measured by the cross-correlation of the speech envelope with the recorded EEG. However, in Mirkovic's study, the simulated spatial experimental setup did not allow for individual differences in head geometry, potentially causing altered externalization of the presented sounds, while the presentation over insert-earphones cannot mimic the changes in level and timing differences normally associated with head movements.

The aim of the present study is to investigate the effect of directional sound processing in a hearing-aid on the neural speech tracking recorded in older hearing-impaired listeners in a truly spatial experimental setup. In the experiment, amplification and directional sound processing was provided by wearable hearing aids.

Hearing aids are able to perform directional sound processing, hereby denoted directionality, whereby it is possible to suppress sounds from different spatial positions. As each hearing aid is equipped with two microphones, the input of these can be combined with different delays resulting in differential sensitivity to sounds from different locations, usually with the purpose of improving the SNR of sounds originating from the frontal plane. In some studies, the strongest possible attenuation profile, e.g. cardioid directivity pattern, is applied to provide the best possible attenuation of the background noise and thus improved the experimental

contrasts. However, in practice, hearing aids often apply a conservative adaptive procedure whereby the directionality automatically adjusts to attenuate sounds from the directions estimated to be noise sources, while at the same time avoid attenuating sounds from any other directions to allow for better speech intelligibility and general scene awareness. Practically, this means that the most aggressive attenuation pattern of noises coming from the back (cardioid with a single null at 180 degrees) is only obtained in situations with a high degree of noise originating from diffuse spatial locations. To avoid artificially maximizing the experimental effect of the applied directionality, the current study applies the directional pattern which the hearing aids would naturally adapt to in the experimental setup.

By applying the backward approach to the Temporal Response Function (TRF) method to the recorded EEG, it is possible to reconstruct the envelope of the to-be-attended speech (Crosse et al., 2016). In the current study a single backwards model is constructed based on speech presented in quiet, which is applied to the EEG recorded from conditions with and without the presence of background noise to reconstruct the speech envelope. Through cross-correlation between the envelope reconstructed by the backward model and the actual speech envelope, it is quantified how similar the neural speech tracking is in quiet and in the presence of background noise. In a similar manner, the same model-reconstructed envelope is correlated with the envelope of the background noise in order to investigate to what degree the background noise is tracked in the same way as speech in quiet. It is hypothesized that hearing-aid directionality (1) improves the resemblance between speech in quiet and speech presented in noise, (2) while the similarity between the background noise and reconstructed speech will be reduced. The forward implementation of the TRF method, will be used to model the neural response to the ongoing speech, hereby denoted the *neural encoding*. It is hypothesized that applying directionality will cause a faster neural encoding response with larger magnitudes, similar to findings reported by Mirkovic et al. (2019). It is further expected that when listening to speech in quiet, the neural encoding response will be faster and have a larger magnitude than the encoding response resulting from listening in noise when directionality is applied.

Methods

Participants

The aim was to include 15 participants in the study, inspired by Mirkovic et al. (2019), but due to COVID-19 restrictions, only 13 participants were tested before lock-down, from which 11 EEG datasets were recorded. All 13 of the older hearing-impaired participants, 5 females, were native Danish speakers. The participants had a mean age of 63.0 years ($sd=5.5$) ranging from 56 to 71 years. They all

suffered from mild-to-moderate sensorineural hearing loss, see Figure 1B, with no more than 15 dB HL difference between the ears. The average pure-tone average (PTA_{8k}, across 0.5, 1k, 2k, 4k and 8 kHz) between ears were 54.8 dB HL ($sd=12.7$) and ranged from 29.0–72.0 dB HL. The average PTA_{8k} across ears was not significantly correlated with age ($r=-0.11$, $p=0.71$). All participants had been wearing hearing aids for more than a year (experienced users).

All participants gave their informed consent and were not given financial compensation for their participation. The study was approved by the by the regional ethical committee of the Capital Region of Copenhagen, Denmark (H-20068621).

Hearing-Aid Processing

All participants were fitted with MOMENT behind-the-ear hearing aids (Widex A/S, Denmark) with individual amplification according to their hearing-threshold levels. To avoid direct-sound inputs, closed fittings were used, with the closeness verified by negligible leakage of sounds presented from the hearing-aid receivers. Finally, the participants audiogram was individually adjusted by measuring in-situ hearing thresholds at frequencies 0.5, 1, 2, and 4 kHz, with no further fine-tuning made.

Two programs were added to the hearing aids, both with sound-classification and noise-reduction algorithms disabled. One program incorporated stationary omni-directional amplification, while the other program applied a fixed directional pattern of the two monaural directional systems (one for each hearing aid). To avoid adaptation of the directional program during the EEG trials, a directional pattern was logged from the hearing aids after adaptation and implemented in a program with fixed, but realistic, directionality. The applied directional pattern was logged from hearing aids on a KEMAR mannequin (Knowles' Electronics Manikin for Acoustic Research, GRAS acoustics, Denmark) placed in the position of the participant and set to the standard adaptive directional program. The stimuli used in the EEG experiment was presented for 45 s, allowing for adaptation of the directionality, before the parameters for the directional pattern were logged from the hearing aids. The final realistic fixed directional pattern applied in the EEG experiment was determined by averaging the parameters across 15 of such measurements. The applied directional pattern, see Figure 1A, shows around 5 dB attenuation across frequencies for the background noise presented from the three loudspeakers from the back, while sounds from the front and sides are left unattenuated.

Sound Material

The to-be-attended stimuli of the experimental task were clips from a Danish version of 'Robinson Crusoe' by Daniel Dafoe narrated by a male speaker. This was chosen because of the monotone reading style (few changes in the

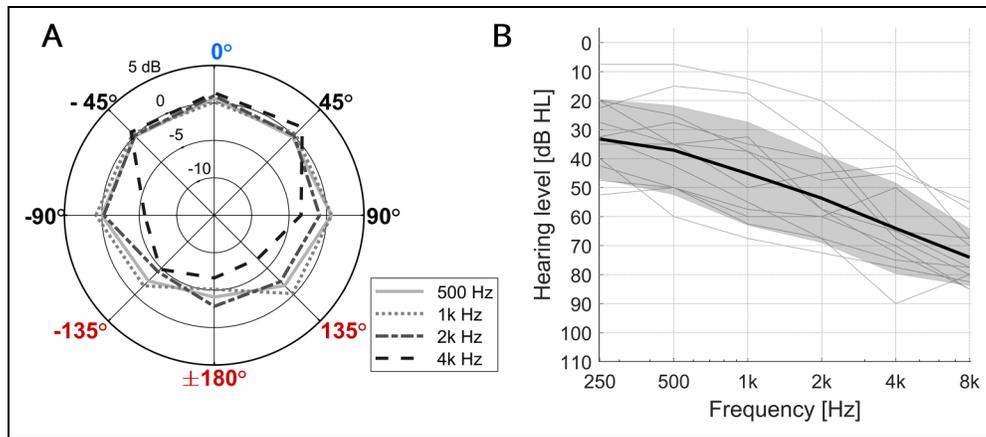


Figure 1. Hearing-aid directionality and participants audiograms. A: Directional pattern of the hearing aids. The lines indicate the attenuation, in dB, of different frequencies of sound from different azimuth directions. Azimuth angles highlighted in red indicate the position of the background noise in the EEG experiment, the angle in blue indicates the position of the target speech. Note that the attenuations depicted for the negative azimuth angles were recorded from the hearing aid positioned at the left ear of the KEMAR, while those at the positive azimuth angles were recorded from the right hearing aid. B: Mean hearing-threshold levels across ears and participants (bold black line) and the standard deviation (shaded grey area). Individual audiograms across ears are shown in thin grey lines.

speech level) and the good sound quality. Using the open-source software Audacity (<https://www.audacityteam.org>), the audiobook was pre-processed by truncating silent periods, with amplitudes below -35 dB, longer than 250 ms down to 180 ms and saved in shorter clips of minimum 64 s at a sampling frequency of 44100 Hz. Since the original narration contained potentially offensive wordings, some clips were edited to not including these terms.

The background noise was a canteen scenario recorded using a single microphone from a lunchbreak at Danish headquarter of WS Audiology A/S. To avoid phase-aliasing, the recording was cut to start at different times when presented from the three loudspeakers behind the participants, see Figure 1A. The spectrum of the canteen noise was not matched to that of the target speaker.

Individual SNR Adjustment

As both hearing loss and working memory capacity affect the ability to understand speech in the presence of background noise (Lunner et al., 2009), the speech was presented at individualized SNRs to avoid large variations in speech intelligibility across participants. The background noise levels were presented at a fixed level of 68 dB-A SPL (sound level meter, type 2250, Brüel & Kjaer, Denmark), while the level of the to-be-attended speech was varied to obtain a speech understanding of 50% (Speech-Reception Threshold of 50% / SRT50) for the individual.

Prior to the experimental task, each participant listened to and repeated 20 sentences from the Danish version of the Hearing-In-Noise Test (HINT; Nielsen & Dau, 2011). The spatial setup of the HINT test was identical to that used in the experimental task, see Figure 1A, but with the sentences

spoken by a different male speaker. The participants were instructed to repeat the sentences to the best of their ability, and the correctness of the sentences were scored by the test leader. The level of the speaker was adjusted in a step-wise approach to obtain the SRT50 (Levitt, 1970). During the HINT test, participants were wearing the experimental hearing aids put in the omni-directional program.

Experimental Task

The participants were seated in the middle of a semi-reverberant 3 by 4-meter sound-proof booth and were instructed to always focus on the voice of the male speaker presented from a loudspeaker 1.5 meter in front of them (0° azimuth, marked with blue in Figure 1A). The narration was disturbed by canteen noise presented at 68 dB-A SPL from three loudspeakers behind the participant ($\pm 135^\circ$ and 180° azimuth, marked with red in Figure 1A). In order to examine the neural effect of directional hearing-aid processing, the scenario was tested with the hearing aids in the omni-directional program and in the directional program. These two conditions are denoted OMNI and DIR, respectively.

In all trials, the male speaker was presented in quiet for 2 to 3 s (time intervals randomly jittered) before the background noise was turned on and presented together with the speech signal for 60 s. The 2 to 3 s baseline period was not included in the analysis of the EEG. Data were also obtained from an additional experimental condition where no background noise was presented, denoted CLEAR. In the CLEAR condition, the target speech continued in silence after the initial baseline period.

Immediately following the 60 s stimulation period, all sounds were turned off and the touchscreen in front of the

participants displayed a three-alternative forced-choice question related to the content of the clip just presented. The question related to either the first or last 30 s of the narration which was chosen at random throughout the experiment. After the participant answered, a red cross appeared on the screen for 2.5 to 4.5 s (time interval randomly jittered) to indicate that a new trial was about to begin. At the onset of the target speech, the cross turned black and remained on the screen during the sound presentation. Following every fifth trial, the participants were prompted to take a pause of self-administered length.

Twenty trials were presented for each of the OMNI and DIR conditions, while the CLEAR condition was presented 15 times. The conditions were presented in a pseudo-randomized order, ensuring that each condition was presented at least once in each of the 9 five-trial blocks. Prior to the presentation of these 45 trials, three training trials were presented in the order CLEAR, OMNI, and DIR. The training trials were not included into the final analysis.

The audiobook clips were presented in a pseudo-randomized order. Following initial pilot experiments (data not included here) with complete randomization of the clips, participants commented that it was difficult to stay engaged when they were not able to follow the flow of the story. Therefore, the clips were randomized within each block, such that clips 1-3 were presented randomly in the training, clips 4-8 were presented randomly in the first block, and so on.

In order to obtain a subjective measure of listening effort, the participants were prompted with the question “How effortful was it for you to follow the story presented from the loudspeaker in front of you?” (translated from Danish). The participants were to indicate their response on a 100-point scale from 1, marked “not effortful”, to 10, marked “extremely effortful”. For time considerations, participants were only asked for this evaluation five times for each of the three condition. The subjective evaluations were performed in a pseudo-randomized fashion such that each block contained at least one, but no more than two, subjective ratings.

EEG Recording and Preprocessing

During the experimental task, neural activity was recorded using the g.tec USBamp (g.tec, Austria). After the HINT test, the cap and corresponding 16 passive electrodes were mounted at standard 10-20 positions equally distributed across the scalp (FP1, FP2, F3, Fz, F4, C3, Cz, C4, T8, T7, P3, Pz, P4, PO1, PO2, and Oz) with a ground electrode at the forehead. Linked mastoids were used as the reference electrode, and the EEG was sampled at 512 Hz. The impedance of all electrodes was kept under 50 kOhm. The EEG was recorded using the g.Recorder software and was co-registered with trigger signals indicating the onset of the to-be-attended speech, the onset of the background noise, and the offset of all sounds.

Due to an incompatibility between g.Recorder software and Windows 10, two datasets were lost, resulting in a total of 11 EEG datasets. Furthermore, this incompatibility caused missing/unregistered triggers, resulting in only 3 of the 11 acquired datasets containing all trials. For the three conditions, an average of 19.0, 19.2, and 14.7 trials were registered for the OMNI, DIR, and CLEAR, respectively. Three independent equivalence tests were conducted using two one-sided t-tests (Schuirmann, 1987), the tests confirmed that the percentage of missing trials did not differ more than 10% (equivalence interval with lower bound -10% and upper bound $+10\%$) between any two conditions (all p 's < 0.05). From the co-registration of the triggers by the stimulus-presentation software (Matlab R2018b), it was possible to confirm that the triggers in the EEG recordings were correctly timed.

All EEG data were analyzed using customized Matlab scripts (R2018b, MA, USA) and the Fieldtrip toolbox (Oostenveld et al., 2011). Trials were extracted from -10 to 70 s relative to the onset of the background noise. All trials were then bandpass filtered between 1 and 45 Hz (3^{rd} order Butterworth) and further notch-filtered between 47 and 53 Hz (4^{th} order Butterworth) to fully suppress the power-line noise. The filters were applied using the Matlab function *filtfilt*, resulting in zero phase distortion. After filtering, the trials cut between -4.5 and 64 s relative to the onset of background noise. All data were visually inspected in order to identify bad channels suitable for interpolation, however none were identified.

An Independent Component Analysis (ICA) analysis was applied for artefact removal. Components related to eye blinks and movements were identified subjectively based on inspection of the time course and topographic distribution and rejected. On average 2.2 of the 15 components of each recording were removed (13.6% of the components, range: 1 to 4 components removed).

Calculating Neural Speech Tracking

The speech envelopes were extracted from the presented audiobook clips by computing the absolute value of the Hilbert transform and low-pass filtering this at 25 Hz (5^{th} order Butterworth, Matlab *filtfilt* function) after which the signal was down-sampled to 64 Hz using the Matlab function *resample* applying an anti-aliasing FIR lowpass-filter. A similar approach was taken to extract the envelope of the overall background noise, in order to evaluate the reconstruction accuracy of the to-be-ignored part of the auditory scene. The EEG was similarly low-pass filtered at 25 Hz and down-sampled to 64 Hz.

It is important to note that the backward modeling approach was implemented by training one subject-specific model on data only from the CLEAR condition. The trained model was then applied to reconstruct the speech envelopes from the OMNI and DIR conditions by providing the model

with EEG from these conditions. The approach of training the model on data from the CLEAR model was taken to provide a common point-of-reference for interpreting the speech reconstruction accuracy, as individualized SNR values were used during the experiment and condition-specific models potentially account for differences between the conditions. Hence, the resulting reconstruction accuracies indicates how similar the reconstruction of the speech presented in the OMNI and DIR condition is to the reconstruction of speech presented in quiet (CLEAR). The reconstruction accuracies of the background noise, obtained by correlating the reconstructed speech envelope with the envelope of the background noise, should also be interpreted as how much the reconstruction of the background noise resembles the encoding of speech in the CLEAR condition. The underlying hypothesis behind this approach to investigate the neural tracking of the background noise is that if the background noise is not fully suppressed during the auditory processing, it will be neural tracked similarly to speech. All computations were done using the mTRF toolbox (version 2.0, Crosse et al., 2016).

To avoid overfitting the backward TRF model generated for each participant trained on the data of the CLEAR condition, a nested leave-one-out cross-validation was implemented (Varma & Simon, 2006). The approach optimizes the individual ridge-regression parameters, λ , in an inner loop, while evaluating the final model on test data in an outer loop, as depicted in Supplemental Figure S1 of the Supplement. A single test-trial is held in the outer loop, while the remaining N-1 trials are passed into the inner loop. In the inner loop another single trial is held back while k backwards TRF models are trained on the remaining N-2 trials using different values of the ridge-regression parameter, $\lambda_k = 10^k$, with $k = \{-6, -4, -2, 0, 2, 4, 6\}$. The performance of the resulting k models, were then evaluated against the single held-back trial of the inner loop in order to determine the individually optimal λ based on the highest Pearson's correlation between the reconstructed and the true speech envelope, denoted the reconstruction accuracy. A final model was then trained on the N-1 trials using the chosen λ and applied to the single test trial of the outer loop to obtain the reconstruction accuracy of that trial. This approach was then repeated for all N trials of the CLEAR condition. As a control measure, the reconstructed speech envelopes from the CLEAR condition were correlated with a randomly chosen speech envelope, not presented in the trial that the model-input EEG originated from. This control condition will be denoted as the random condition. Note that the number of trials in the CLEAR differed between individuals (12 to 15 trials, mean 14.7 trials) due to incorrect triggering of the EEG recording (see EEG Recording and Preprocessing). All model was specified to include EEG within the time-lags from -100 to 700 ms relative to the onset of the speech.

To generate a reconstructed envelope for each trial of the OMNI and DIR conditions, a single final CLEAR model was

obtained for each participant by averaging across all models of the outer loops and applied to the EEG of the OMNI and DIR conditions. For each trial, the reconstructed envelope was also correlated with the actual presented speech envelope as well as the envelope of the background noise. In the analysis of the background noise, the accuracies arising from the CLEAR condition will act as a control measure, as no background noise was presented in this condition. The resulting correlation values from the OMNI and DIR conditions express how similar the neural tracking of the background noise is to the tracking of speech in the absence of background noise (CLEAR).

The neural encoding of the speech was investigated through the modelled neural response, obtained by training forward TRF models for each of the three experimental conditions within the time-lag -100 to 700 ms. For each condition and participant, the ridge-regression parameter λ was chosen from the subset, $\lambda_k = 10^k$ with $k = \{-6, -4, -2, 0, 2, 4, 6\}$, in an outer- and inner-loop implementation as described above. The corresponding model resulting in the highest Pearson's correlation between reconstructed and recorded EEG was applied to generate the modelled neural encoding response. Again, a control response was generated by training a model with EEG from the CLEAR condition and randomly chosen speech envelopes.

Statistical Analysis

Statistical differences in behavioral (performance and listening-effort ratings) and neural (speech reconstructions accuracy) measures averaged across trials for each participant were investigated using linear mixed-effects regression (LMER) models implemented in the *lme4* package for R (Bates et al., 2015). All models included fixed effects of condition and PTA_{gk} , the interaction between them and the random effect of participants. The degree of hearing loss was included in the analysis to account for potential effects thereof, see Introduction for further elaboration. The residuals of all the LMER models did not significantly deviate from a normal distribution, confirmed by non-significant Shapiro-Wilk tests. All post hoc analysis of statistically significant effects were performed using the *lm_means* package (Kuznetsova et al., 2017). The guidelines proposed by Althouse (2016) will be followed by specifying the effect size (regression coefficient), confidence interval, and p-value for all reported effects (Althouse, 2016).

Comparing the time-varying neural encoding responses (forward TRF modelled neural responses) between conditions requires adequately control for the multiple comparisons made across time-lags and electrodes. For this purpose, a two-level cluster-based approach implemented in the Fieldtrip toolbox was applied (Maris & Oostenveld, 2007). In short, this approach first identifies time-channel combinations showing condition effects for each participant and on the second level, it tests whether the identified combinations are different

from zero across participants. On the first (participant-)level, the time-channel neural encoding response of the three conditions were compared using an independent-samples regression analysis by assigning each condition a linearly spaced weight of -1 , 0 , or $+1$ for OMNI, DIR and CLEAR, respectively. These zero-centered weights specify that a similar change (positive or negative) is expected between OMNI and DIR and between DIR and CLEAR. The resulting t -values were transformed into z -scores. On the second (group-)level, the individual z -scores were tested against zero using a two-sided dependent t -test. From the corresponding p -values, clusters were formed based on adjacent time and channels. A minimum of three adjacent channels had to show a significant response with $p < 0.01$ to form a cluster and the corresponding test statistics were summed across the t -values within the cluster. For the final determination of statistical significance, the t -value of each cluster was compared to a distribution of t -values generated by producing 5000 permutations of randomly assigning the condition labels (zero or z -score) and calculating the corresponding t -value (for details, see Maris & Oostenveld, 2007).

Results

Task Performance and Subjective Listening Effort

The HINT test conducted prior to the EEG experiment with participants being aided (OMNI) resulted in SRT50-value ranging from -2.7 to 9.3 dB SNR (mean = 3.2 , sd = 4.0 dB SNR). A Pearson's correlation showed that listeners with worse hearing (PTA_{8k}) required significantly higher SNR values to understand 50% of the sentences [$r = 0.8$, CI: 0.42 – 0.93 , $p < 0.01$].

Applying the individualized SRT50-values from the HINT test in the EEG experiment, resulted in the percentage of correctly answered three-alternative forced-choice questions being significantly higher than the chance level of 33.3% [Figure 2A, $t_{38} = 24.2$, CI: 0.45 – 0.53 , $p < 0.01$]. A linear mixed-effects model revealed that the performance was significantly affected by conditions [$F_{2,22} = 11.9$, $p < 0.001$]. The post hoc analysis revealed that OMNI resulted in poorer performance than DIR [$t_{22} = 4.13$, $r = 0.13$, CI: 0.07 – 0.21 , $p < 0.001$] and CLEAR [$t_{22} = -4.32$, $r = 0.15$, CI: 0.07 – 0.22 , $p < 0.001$], but did not reveal any difference between DIR and CLEAR [$t_{22} = 0.19$, $r = 0.01$, CI: -0.06 – 0.08 , $p = 0.84$].

The subjective ratings of listening effort were also affected by the experimental conditions [Figure 2B, $F_{2,22} = 7.3$, $p < 0.001$], with lower effort ratings in CLEAR compared to OMNI [$t_{22} = 3.7$, $r = 2.4$, CI: 1.1 – 3.8 , $p < 0.001$] and DIR [$t_{22} = 2.5$, $r = 1.7$, CI: 0.3 – 2.9 , $p = 0.02$]. However, the participants did not rate the listening effort as different between DIR and OMNI [$t_{22} = 1.2$, $r = 0.7$, CI: -0.6 – 2.2 , $p = 0.24$].

From the analysis of the behavioral outcomes, the performance improved significantly with the addition of directionality (OMNI vs. DIR), whereas no significant differences were found in the subjective rating of listening effort between OMNI versus DIR.

No significant effects of PTA_{8k} was found on the task performance or the ratings of listening effort [all p 's > 0.2].

Effect of Directionality on Neural Speech Tracking

Figure 3 shows the reconstruction accuracies, i.e. the Pearson's correlation between the speech reconstructed using a model based on data from the CLEAR condition and the actual envelope of the to-be-attended speech (Figure 3A) and the to-be-ignored background noise (Figure 3B) for all listening conditions. For the reconstruction of the background noise, the CLEAR condition can be considered a control, similar to the random condition for the reconstruction of the speech signal.

A significant effect of condition was found on the reconstruction accuracy of the to-be-attended speech [$F_{3,27} = 59.2$, $p < 0.001$]. The post hoc analysis revealed that all three experimental conditions (CLEAR, DIR, OMNI) had significantly higher reconstruction accuracies than the reconstruction of a random speech signals (all p 's < 0.001). For the CLEAR condition, a significantly better reconstruction of the speech envelope was obtained compared to both conditions with background noise [DIR: $t_{27} = 2.2$, $r = 19 \cdot 10^{-3}$, CI: 2 – $36 \cdot 10^{-3}$, $p = 0.03$. OMNI: $t_{27} = 4.6$, $r = 39 \cdot 10^{-3}$, CI: 22 – $56 \cdot 10^{-3}$, $p < 0.001$]. Applying directionality also significantly improved the speech reconstruction (OMNI vs. DIR, $t_{27} = 2.4$, $r = 20 \cdot 10^{-3}$, CI: 3 – $37 \cdot 10^{-3}$, $p = 0.02$). No effects of PTA_{8k} was identified (all p 's > 0.3).

An alternative to expressing the reconstruction accuracy as the cross-correlation coefficient, the percentage of explained variance ($R^2 \cdot 100$) can be examined. For the reconstruction of the speech, the percentage of variance explained (average across subjects for OMNI = 0.5% , DIR = 0.8% , CLEAR = 1.2%) showed the same significant differences between conditions as the analysis of the correlation coefficients.

The similarity between the envelope of the background noise and that of the reconstructed attended speech (Figure 3B) showed no significant effect of condition [$F_{2,18} = 1.3$, $p = 0.3$] or hearing loss [$F_{1,9} = 0.4$, $p = 0.5$]. The Bayesian Information Criterion (BIC) of this model proved to be larger than the BIC of a model only containing a constant term (corresponding to a backwards elimination of all non-significant factors), confirming that the envelope of the background noise cannot be reliably reconstructed using the applied approach.

Task Performance in Relation to Neural Speech Tracking

A further area to explore is the relation between speech reconstruction accuracies and task performance. The speech

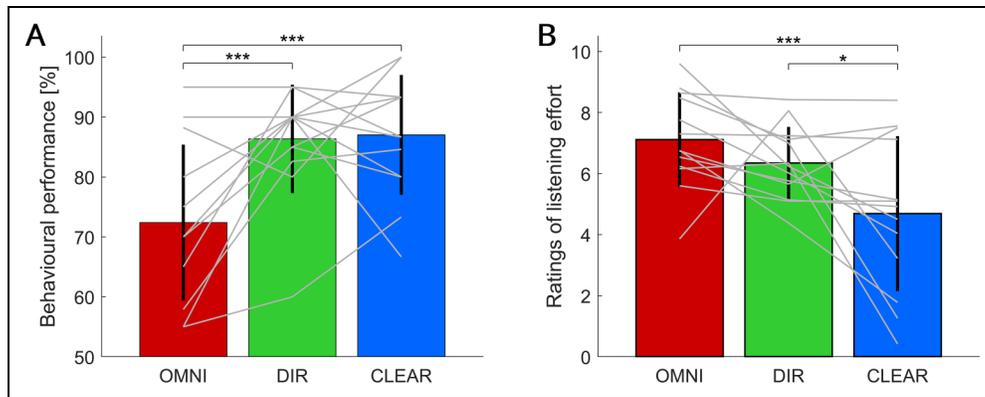


Figure 2. Performance and listening effort ratings. A: Percentage of correctly answered control questions for the three experimental conditions. Note that the y-axis starts at 50%, well above the chance level at 33.3% correct. B: Subjective ratings of the listening effort averaged across the five evaluations made for each condition per participant. The height of each bar denotes the mean across participants ($n = 13$), and the black line indicates ± 1 sd. Individual results are shown in grey lines. Asterisks indicate the p-value of the statistically significant differences between conditions (* $p < 0.05$; *** $p < 0.001$).

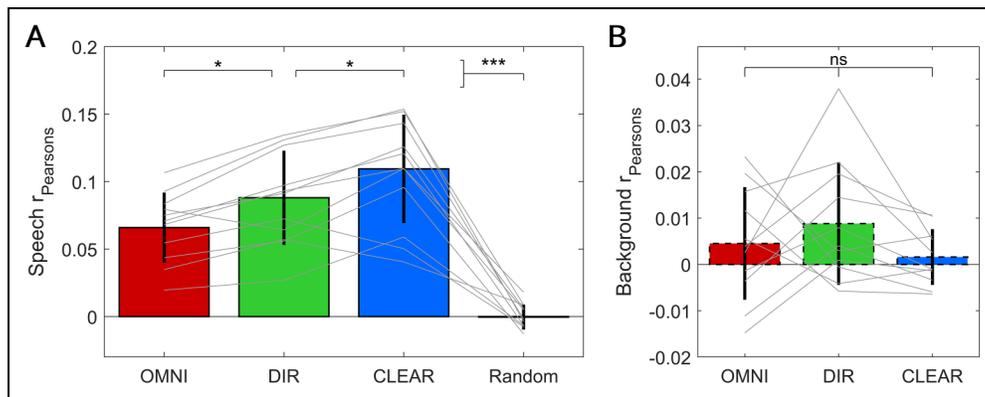


Figure 3. Neural speech tracking results. A: The speech envelope reconstruction accuracies, as measured by the Pearson's correlational values, of the to-be-attended speech in all three listening conditions. The random condition represents the correlation between a reconstructed speech envelope from the CLEAR condition and a randomly chosen speech envelope. B: Similarity between the reconstructed speech envelope and the envelope of the to-be-ignored background noise, with the CLEAR condition corresponding to a control condition as no noise was present in that condition. The height of each bar denotes the mean across participants ($n = 11$), and the black line indicates ± 1 sd. Individual results are shown in grey lines. Asterisks indicate the p-value of the statistically significant differences between conditions (* $p < 0.05$; *** $p < 0.001$, ns: non-significant).

accuracies of the random control condition were omitted, and the final model included condition (CLEAR, OMNI, DIR), the task performance (correct or incorrect) of each trial, and the interaction between these two factors as fixed effects, participants were included as a random effect. Besides the expected effect of condition [$F_{1,566.3} = 18.2$, $p < 0.01$], an effect of trial performance was observed [$F_{1,568.2} = 4.9$, $p < 0.01$], revealing that the reconstruction of the to-be-attended speech was significantly better for the correctly answered trials than in trials with incorrect answers [$t_{566.3} = 2.2$, $r = 11 \cdot 10^{-3}$, CI: $1-21 \cdot 10^{-3}$, $p = 0.03$].

The subset of trials where the listening effort was evaluated was extracted, and a model similar to the above

described applied, but no relationship between ratings and the corresponding reconstruction of the to-be-attended speech was found [$F_{1,154} = 0.07$, $p = 0.8$] of the same trials.

Finally, when substituting PTA_{8k} by the individually adjusted SNR, no statistical effect thereof was found on the reconstruction accuracy of the speech.

Neural Encoding of To-Be-Attended Speech

The individually modelled neural responses arising from the weights of the forward TRF model were explored to assess the effect of noise and hearing-aid directionality on the brain's neural encoding of speech, see Figure 4. The cluster-

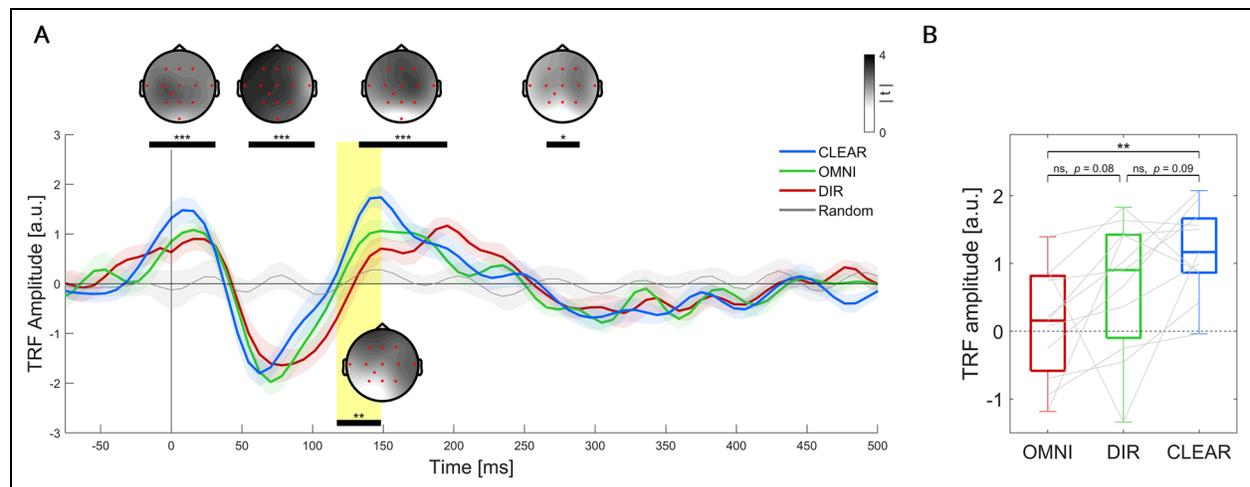


Figure 4. Neural encoding results. A: The time course of the TRF forward model weights for each experimental condition and the random condition. The solid lines show the model weights averaged across models for each participant ($N = 11$) and 16 EEG channels with the shaded areas indicating the 95% confidence interval of each response across participants. Above the graph: The bold black lines and topographic maps indicate the time and electrode range of clusters where responses of the experimental conditions differ significantly from the random condition. The red dot in the topographic map indicates the significant electrodes, while the shading shows the absolute t -values of the cluster-based statistics. Under the graph: The black line, yellow shaded, and topographic map indicate the interval where the model responses were significantly affected by the experimental conditions. B: Model weights for each individual and experimental condition when averaged within the cluster indicating a significant difference between conditions (yellow area in A). The central line of each box indicates the median, the edges of the box the 25th and 75th percentile, the whiskers the 95% confidence interval. The individual results are plotted in grey lines. Asterisks indicate the p -values of the statistical test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns: non-significant).

based analysis revealed that the neural encoding responses were different from the random model in four time intervals, see Figure 4A: A complex of peaks consisting of a positive [-16 to 31 ms, 13 electrodes, $p < 0.001$], a negative [55 to 102 ms, 13 electrodes, $p < 0.001$], and a positive peak [133 to 195 ms, 13 electrodes, $p < 0.001$]. Additionally, a fourth negative deflection was detected after 250 ms [266 to 289 ms, 12 electrodes, $p = 0.015$].

The two-level statistical approach revealed a single cluster in which the model parameters were significantly affected by the experimental conditions [117 to 149 ms, 12 electrodes, $p < 0.01$, yellow area in Figure 4A], spanning the transition between the first negative and second positive deflections in the modelled neural encoding responses. The model weights for each participant and experimental condition were averaged across the significant time and channels, see Figure 4B, and further investigated in an LMER analysis. The model confirmed a significant effect of experimental condition [$F_{2,18} = 6.5$, $p < 0.01$], while no significant effects of hearing loss was found (all p 's > 0.5). The post hoc analysis of the condition effect showed that OMNI had significantly lower model weights within the cluster compared to CLEAR [$t_{18} = 3.6$, $r = 1.07$, CI: 1.7–3.6, $p < 0.01$]. The average model weights for DIR did not significantly differ from those of OMNI [$t_{18} = 1.8$, $r = 0.5$, CI: -0.09–1.1, $p = 0.08$] and CLEAR [$t_{18} = 1.7$, $r = -0.54$, CI: -0.09–1.1, $p = 0.09$], however indicated a trend of decreasing model response weights as noise was added and directionality removed.

Discussion

In this study, the effect of hearing-aid directionality on the neural speech tracking and neural encoding of speech was investigated. Our main findings were that: (1) Task performance improved when applying directionality (DIR vs. OMNI), although this was not reflected in the subjective ratings of listening effort. (2) The neural tracking of the to-be-attended speech was significantly improved with directionality (DIR vs. OMNI) and when the background noise was removed (DIR vs. CLEAR). However, (3) the neural tracking of the background noise did not statistically differ between the experimental conditions. Finally, (4) the neural encoding of speech became gradually faster and larger when directionality was added and background noise removed. In the following, the findings will be discussed in further detail.

Directionality Affects Performance, but Not Listening Effort

At the end of each 1-min trial, the participants answered a three-alternative forced-choice control question relating to the content of the clip just presented. The increase in performance resulting from applying directionality confirms that directionality improves speech intelligibility, see Figure 2A. It is interesting that it cannot be concluded from the data that removing the background noise (CLEAR vs. DIR),

resulted in higher behavioral task performance. It is possible that a performance ceiling was reached for DIR and CLEAR and that the difficulty of the control questions result in this level being around 85% correct. However, it might also be speculated that listening to speech in quiet (CLEAR) simply is so easy and unengaged (Herrmann & Johnsrude, 2020), that participants do not hear the answer to the control question, e.g. due to mind wandering (Varao-Sousa et al., 2018).

Participants rated that their listening effort was lowest in the condition without background noise (CLEAR), but no significant difference between OMNI and DIR was found (see Figure 2B). Thus, it is possible that the SNR difference between OMNI and DIR was not large enough to trigger a difference in self-perceived listening effort. However, as the term ‘*listening effort*’ is not a well-defined concept, participants might include a range of individual factors, such as the general level of fatigue or stress, in their subjective evaluation of listening effort (McGarrigle et al., 2014; Pichora-Fuller et al., 2016).

All things considered, it is interesting to observe that the statistical analysis revealed that the task performance and rated effort were significantly affected by two different factors, adding directionality and adding noise, respectively. The behavioral outcomes of the current study provide a great example of a situation where neural outcome measures could provide a more nuanced picture of the effect of hearing-aid directionality.

Neural Speech Tracking Improves with Directionality

When applying a backward implementation of the TRF model to reconstruct the to-be-attended speech, it was observed that the similarity between pure speech (CLEAR) improved significantly with directionality (OMNI vs. DIR), but also when removing the background noise compared to attenuating it (DIR vs. CLEAR). The condition effects for the individual participants in Figure 3A, show that the increase in reconstruction accuracy between OMNI, DIR, and CLEAR was observed for 9 of the 11 participants. This very clearly shows that while the presence of background noise reduces neural speech tracking, applying hearing-aid directionality further improves it.

A similar result was reported by Alickovic et al. (2020) finding that the to-be-attended speech was better reconstructed when applying a combination of both directional sound processing and noise reduction. In the same study, an improvement in speech tracking of the attended speech was not observed when the SNR was reduced from +8 to +3 dB SNR, an alteration approximately corresponding to the effect of the applied signal processing scheme. This observation could suggest that the noise-reduction, rather than the directionality, was driving the effect of signal processing observed by Alickovic and colleagues. In the current study however, we did observe an effect of reducing

the SNR using directionality alone when training the model on the CLEAR data, Figure 3A, as well as when training the model on data from all the experimental conditions, see Supplemental Figure S2.

It should be kept in mind that the all speech reconstruction results are based on a model trained on speech from the CLEAR condition. Hence, the neural speech tracking of the OMNI and DIR conditions reflects how well the output of the backwards model resembles speech presented in the absence of background noise (CLEAR). A more traditional approach is to train a model for each experimental condition, however with the caveat that the models inherently change between conditions, thereby making the effect of the experimental conditions difficult to interpret. Opposed to studies applying the TRF approach for optimizing speech reconstruction accuracies to be applied in attention-controlled hearing devices (for overview, see Geirnaert et al., 2021), the goal of the current study was not to maximize the neural speech tracking, but to focus on evaluating the experimental contrasts on an easily-interpretable basis. Alternatively, training a model on all data reflects how well the speech is neural encoded relative to any kind of speech (embedded in noise or not), which is more suitable for drawing conclusions on the generalizability of the results. Although it is not the scope of this study, an investigation was made looking into whether the condition effects persisted when training model on different data (see Supplemental Figure S2). The results revealed that significant differences in reconstruction accuracy between OMNI and DIR persisted when training the backward TRF models on 1) data from the DIR condition only, 2) data from each condition individually and 3) on data from all three experimental conditions (see Figure S1 in Supplement). However, when training a model on data from the OMNI conditions, no significant differences between the reconstruction accuracies were seen between the reconstruction accuracies of the different experimental conditions.

From the reconstruction accuracies presented in the Supplement, it is also evident that the improvement in accuracy observed between DIR and CLEAR is not present for any of the other model (Supplemental Figure S2). This suggests that basing the model on CLEAR data can highlight the subtle difference in neural speech tracking between CLEAR and DIR, likely caused by the CLEAR model not factoring in the effect of adding background noise. That the general encoding of speech is affected by the presence of noise is confirmed by the neural encoding responses resulting from the different conditions (Figure 4), showing that the reconstruction of speech presented in background noise is neurally tracked and processed differently from speech presented in quiet (OMNI vs. CLEAR).

The similarity between the reconstructed speech envelope and the envelope of the background noise did not show to be different for the OMNI and DIR compared to the control condition (CLEAR, see Figure 3B), indicating that the applied

approach was not suitable for quantifying the disturbance of the background noise. This could of course be due to the lack of asserted attention, known to be reduced for to-be-ignored signals (Ding & Simon, 2012; Power et al., 2012), or the relatively stationary envelope of the babble-like canteen noise used in this study. However, it is possible that the approach expressing how similar the tracking of the background noise is to the tracking of CLEAR speech is inadequate.

Recently, it has been found that comprehension of a target speaker was reduced when magnetically stimulating with the envelope of distracting speech at a phase of 0 degrees, indicating an interplay between to-be-attended and to-be-ignored processing, where the enhanced processing of one causes a reduced processing of the other (Keshavarzi, Varano, & Reichenbach, 2021). Although we observe the reconstruction of the to-be-attended speech to increase when adding directionality and removing background noise, we were not able to identify similarities between the reconstructed speech and the background noise which could indicate that the reduced speech intelligibility, as evident from the poorer task performance, was caused by the contamination of the neural tracking of the to-be-attended speech by the to-be-ignored background noise.

Alterations in neural tracking of attended speech can also indicate changes in the speech intelligibility. It has been shown that the speech reconstruction accuracy follows a sigmoidal psychometric curve as a function of SNR (Vanthornhout et al., 2018), similar to the changes in speech intelligibility seen with varying SNR. However, the observations were made for speech reconstructions models including only the initial 75 ms of the neural response, known not to be affected by attention (O'Sullivan et al., 2019). In the current study, the later neural responses to speech, up to 700 ms, were included in the TRF model, hence also containing attentional-driven neural activity.

Furthermore, neural speech tracking is not only affected by whether attention is asserted, but also by the level of this attention. In a recent study, it was found that the neural speech tracking was more accurate in time intervals with a high level of attention, estimated by the spectral entropy of the EEG, compared to periods with low levels of attention (Lesenfants & Francart, 2020). It could be speculated that the attention level estimated by Lesenfants and Francart is a proxy measure of listening effort. Indeed, it has previously been noted that the amount of allocated attention “*could be used to make inferences about listening effort*” (Pichora-Fuller et al., 2016). However, studies investigating how neural speech tracking is influenced by the degree of listening effort experienced by the listener, cannot prove a direct link between the two: Decruy and colleagues found no significant correlation between neural speech tracking and a total of five measures of behavioral, neural, and self-reported listening effort (Decruy et al., 2020). A study by Müller et al., identified no relationship between the neural encoding of speech and the subjectively perceived listening

effort. However, this study also did not find a relationship between the pupil response and listening effort, a relationship described in many other studies (Müller et al., 2019).

In line with the above studies, the current study also did not reveal a relationship between neural speech tracking and subjective ratings of listening effort. In contrast, it was found that the speech reconstruction accuracies were significantly higher for trials where the control question was answered correctly compared to trials with incorrect answers. However, it should be kept in mind that the number of correct trials in the current study vastly outnumbers the number of incorrect. Furthermore, it is not possible to know whether participants answer incorrectly due to reduced speech intelligibility or decreased attentional levels, e.g. through mind wandering.

The degree of hearing loss was included in all statistical models, however none of the behavioral or neural outcome measures were significantly affected by it. Although this seems in line with the results from Presacco et al. and Goossens et al., it should be noticed that with the limited range of hearing impairment (PTA_{8k} range: 29.0 to 72.0 dB HL) and number of participants (n = 11) means that the study was relative insensitive to this factor.

Directionality Results in Faster Neural Encoding of Speech

The weights of the forward models were constructed to evaluate the modelled neural encoding response of the to-be-attended speech for each experimental condition. The cluster-based statistical analysis revealed neural encoding of the speech stimuli in three time-intervals, see Figure 4A. The consecutive positive-negative-positive peaks of the model responses observed in the current study concurs with the results of previous studies (Crosse et al., 2016; Ding & Simon, 2012; Mirkovic et al., 2019; Petersen et al., 2017). The neural encoding response has previously been described to resemble the classical auditory evoked response (P1-N1-P2) in latency and polarity (Crosse et al., 2016). In the current study, a smaller fourth negative deflection of the neural encoding response was detected between 257 and 289 ms, equivalent to an N2-like response, see Figure 4A. Although this fourth deflection is not often mentioned in descriptions of TRF model response, it is evident from previous studies that the response is present (Figure 4C in Ding and Simon 2012; Figure 4 in Hjortkjaer et al. 2018; Figure 2 in Jaeger et al. 2020; Figure 5 in Mirkovic et al. 2019). As traditionally evoked N2 responses have been linked to the allocation of attention and phonological processing (Tomé et al., 2015), it is not unexpected that an N2-like response of the modelled neural encoding response is observed in the current, as well as in other studies. Indeed, the latencies of the TRF model responses have been compared to the properties of evoked responses, with early peaks reflecting initial low-level processing of

the entire auditory scene, whereas the responses occurring after ~85 ms are associated with top-down processing of individual sound sources based on asserted attention (Alickovic et al., 2021; Lesenfants & Francart, 2020; O'Sullivan et al., 2019).

A significant effect of the experimental conditions was identified around 130 ms (range 117 to 149 ms, see yellow area in Figure 4A), where the TRF responses transition from the first negative peak to the second positive peak (equivalent to N1-P2 for evoked potentials). This finding concurs with the findings by Mirkovic et al. also identifying directionality to affect the neural encoding in the transition between the N1-P2 components (Mirkovic et al., 2019).

The current study and the study by Mirkovic et al. differ in many ways, in that we (1) used the TRF model approach and not the cross-correlation between EEG with the speech stimuli, (2) applied real ear-worn hearing aids allowing for a more natural spatial experience, (3) applied a fixed version of the directional pattern which the hearing aids would naturally adapt to in the given situation, (4) based the analysis on the envelopes calculated from the original sound stimuli and not on the envelopes of the amplified sound stimuli, and finally (5) our findings were based on fewer and only hearing-impaired participants. Despite these, relatively large, differences in the experimental approach, the effect of hearing-aid directionality seems consistent between the two studies.

The results of both the current study and the study of Mirkovic et al. suggest a faster transition between the early bottom-up processing of the auditory scene to the more top-down driven processing of the attended speech. Interpreted in the frame of the Ease of Language Understanding model (ELU, Rönnberg et al., 2013), this might reflect that hearing-aid directionality causes a more accurate match between the auditory input and the expected internal representation of the lexical content, leading to a faster understanding. If such a match is not obtained, the cognitive system will have to rely on slower, and more working-memory driven, semantic and phonological processing, to infer meaning from the to-be-attended speech (Rönnberg et al., 2013). The neural encoding responses in Figure 4A, show that adding noise while not applying directionality result in a slowing down of the neural encoding response, potentially caused by the requirement of additional cognitive processing of the incoming to-be-attended speech.

In summary, the current study provides evidence that hearing-aid directionality improves the neural tracking of to-be-attended speech. The neural encoding of speech showed that when directionality is applied, the transition between the early responses involved in processing the auditory scene and the later top-down driven responses modulated by attention becomes faster. This suggests that applying hearing-aid directionality improves the efficiency of the neural processing of speech resulting in a better

neural representation and consequently an improved speech understanding.

Declaration of Conflicting Interests

The author, Eline Borch Petersen, is employed at WS Audiology, Lyngø, Denmark.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Eline Borch Petersen  <https://orcid.org/0000-0001-9258-5795>

Supplemental Material

Supplemental material for this article is available online.

References

- Alickovic E., Lunner T., Wendt D., Fiedler L., Hietkamp R., Ng E. H. N., & Graversen C. (2020). Neural representation enhanced for speech and reduced for background noise with a hearing aid noise reduction scheme during a selective attention task. *Frontiers in Neuroscience*, *14*(September), 846. <https://doi.org/10.3389/FNINS.2020.00846>
- Alickovic E., Ng E. H. N., Fiedler L., Santurette S., Innes-Brown H., & Graversen C. (2021). Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise. *Frontiers in Neuroscience*, *15*(March). <https://doi.org/10.3389/fnins.2021.636060>
- Althouse A. D. (2016). Adjust for multiple comparisons? It's not that simple. *Annals of Thoracic Surgery*, *101*(5), 1644–1645. <https://doi.org/10.1016/j.athoracsur.2015.11.024>
- Bates D., Mächler M., Bolker B. M., & Walker S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). <https://doi.org/10.18637/jss.v067.i01>
- Bernarding C., & Strauss D. Hannemann J., & Corona-Strauss R. F. I. (2012). Quantification of listening effort correlates in the oscillatory EEG activity: A feasibility study. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 4615–4618. <https://doi.org/10.1109/EMBC.2012.6346995>.
- Bernarding C., & Strauss D. Hannemann J., Seidler R., & Corona-Strauss H., F. I. (2014). Objective assessment of listening effort in the oscillatory EEG: comparison of different hearing aid configurations. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014, 2653–2656. <https://doi.org/10.1109/EMBC.2014.6944168>.
- Bernarding C., Strauss D. J., Hannemann R., Seidler H., & Corona-Strauss F. I. (2017). Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort. *Cognitive Neurodynamics*, *11*(3), 203–215. <https://doi.org/10.1007/s11571-017-9425-5>
- Crosse M. J., Di Liberto G. M., Bednar A., & Lalor E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous

- stimuli. *Frontiers in Human Neuroscience*, 10(November), 1–14. <https://doi.org/10.3389/fnhum.2016.00604>
- Dai L., Best V., & Shinn-Cunningham B. G. (2018). Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention. *Proceedings of the National Academy of Sciences*, 115(14), 201721226. <https://doi.org/10.1073/pnas.1721226115>.
- Decruy L., Lesenfants D., Vanthornhout J., & Francart T. (2020). Top-down modulation of neural envelope tracking: the interplay with behavioral, self-report and neural measures of listening effort. *European Journal of Neuroscience*, 52(5), 3375–3393. <https://doi.org/10.1111/ejn.14753>
- Ding N., & Simon J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, 109(29), 11854–11859. <https://doi.org/10.1073/pnas.1205381109>
- Dubno J. R., Dirks D. D., & Morgan D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, 76(1), 87–96. <https://doi.org/10.1121/1.391011>
- Fiedler L., Seifi Ala T., Graverson C., Alickovic E., Lunner T., & Wendt D. (2021). Hearing aid noise reduction lowers the sustained listening effort during continuous speech in noise—A combined pupillometry and EEG study. *Ear & Hearing*, 42(6), 1–12. <https://doi.org/10.1097/aud.0000000000001050>.
- Fuglsang S. A., Märcher-Rørsted J., Dau T., & Hjortkjaer 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>
- Geirnaert S., Vandecappelle S., Alickovic E., de Cheveigné A., Lalor E., & Meyer B. T., ... Bertrand A. (2021). EEG-based Auditory attention decoding: towards neuro-steered hearing devices. arXiv:2008.04569
- Goossens T., Vercammen C., Wouters J., & van Wieringen A. (2019). The association between hearing impairment and neural envelope encoding at different ages. *Neurobiology of Aging*, 74, 202–212. <https://doi.org/10.1016/j.neurobiolaging.2018.10.008>
- Grady C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, 13(7), 491–505. <https://doi.org/10.1038/nrn3256>
- Grady C., Springer M. V., Hongwanishkul D., McIntosh A. R., & Winocur G. (2006). Age-related changes in brain activity across the adult lifespan. *Journal of Cognitive Neuroscience*, 18(2), 227–241. <https://doi.org/10.1162/jocn.2006.18.2.227>
- Herrmann B., & Johnsrude I. S. (2020). A model of listening engagement (MoLE). *Hearing Research*, 397, 108016. <https://doi.org/10.1016/j.heares.2020.108016>
- Hjortkjaer J., Märcher-Rørsted J., Fuglsang S. A., & Dau T. (2018). Cortical oscillations and entrainment in speech processing during working memory load. *European Journal of Neuroscience*, (January), 51, 1279–1289. <https://doi.org/10.1111/ejn.13855>
- Howard M. F., & Poeppel D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns Depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500–2511. <https://doi.org/10.1152/jn.00251.2010>
- Jaeger M., Mirkovic B., Bleichner M. G., & Debener S. (2020). Decoding the attended speaker from EEG using adaptive evaluation intervals captures fluctuations in attentional listening. *Frontiers in Neuroscience*, 14(June), 603. <https://doi.org/10.3389/fnins.2020.00603>
- Keshavarzi M., Varano E., & Reichenbach T. (2021). Cortical tracking of a background speaker modulates the comprehension of a foreground speech signal. *The Journal of Neuroscience*, 41(23), 5093–5101. <https://doi.org/10.1523/jneurosci.3200-20.2021>
- 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>
- Lesenfants D., & Francart T. (2020). The interplay of top-down focal attention and the cortical tracking of speech. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/s41598-020-63587-3>
- Levitt H. (1970). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2, Pt. 2), 467–477. <https://doi.org/10.1121/1.1912375>
- Lunner T., Rudner M., & Rönnerberg J. (2009). Cognition and hearing aids. *Scandinavian Journal of Psychology*, 50(5), 395–403. <https://doi.org/10.1111/j.1467-9450.2009.00742.x>
- Maris E., & Oostenveld R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- McGarrigle R., Munro K. J., Dawes P., Stewart A. J., Moore D. R., Barry J. G., & Amitay S. (2014). Listening effort and fatigue: what exactly are we measuring? A British society of audiology cognition in hearing special interest group “white paper.” *International Journal of Audiology*, 53(7), 433–445. <https://doi.org/10.3109/14992027.2014.890296>
- Mesgarani N., & Chang E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485(7397), 233–236. <https://doi.org/10.1038/nature11020>
- Millman R. E., Mattys S. L., Gouws A. D., & Prendergast G. (2017). Magnified neural envelope coding predicts deficits in speech perception in noise. *The Journal of Neuroscience*, 37(32), 7727–7736. <https://doi.org/10.1523/JNEUROSCI.2722-16.2017>
- 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.jcsr.2012.02.004>
- Müller J. A., Wendt D., Kollmeier B., Debener S., & Brand T. (2019). Effect of speech rate on neural tracking of speech. *Frontiers in Psychology*, 10(MARCH), 1–15. <https://doi.org/10.3389/fpsyg.2019.00449>
- Nielsen J. B., & Dau T. (2011). The danish hearing in noise test. *International Journal of Audiology*, 50(3), 202–208. <https://doi.org/10.3109/14992027.2010.524254>
- O’Sullivan J., Herrero J., Smith E., Schevon C., McKhann G. M., Sheth S. A., ... Mesgarani N. (2019). Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron*, 104(6), 1195–1209.e3.e3. <https://doi.org/10.1016/j.neuron.2019.09.007>
- 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, 156869. <https://doi.org/10.1155/2011/156869>

- 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>
- Pichora-Fuller M. K., Kramer S. E., Eckert M. A., Edwards B., Hornsby B. W. Y., Humes L. E., ... Wingfield A. (2016). Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S–27S. <https://doi.org/10.1097/AUD.0000000000000312>
- Power A. J., Foxe J. J., Forde E. J., Reilly R. B., & Lalor E. C. (2012). At what time is the cocktail party? A late locus of selective attention to natural speech. *European Journal of Neuroscience*, *35*(9), 1497–1503. <https://doi.org/10.1111/j.1460-9568.2012.08060.x>
- 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), 1–26. <https://doi.org/10.1371/journal.pone.0213899>
- Rönngberg J., Lunner T., Zekveld A., Sörqvist P., Danielsson H., Lyxell B., ... Rudner M. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*(31), 1–17. <https://doi.org/10.3389/fnsys.2013.00031>
- Schuirman D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, *15*(6), 657–680. <https://doi.org/10.1007/BF01068419>
- Shinn-Cunningham B. G., & Best V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283–299. <https://doi.org/10.1177/1084713808325306>
- Strauß A., Wöstmann M., & Obleser J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, *8*(May), 1–7. <https://doi.org/10.3389/fnhum.2014.00350>
- Taylor B. (2003). Speech-in-noise tests: how and why to include them in your test battery. *The Hearing Journal*, *56*(1), 40, 42–46. <https://doi.org/10.1097/01.HJ.0000293000.76300.ff>
- Tomé D., Barbosa F., Nowak K., & Marques-Teixeira J. (2015). The development of the N1 and N2 components in auditory oddball paradigms: A systematic review with narrative analysis and suggested normative values. *Journal of Neural Transmission*, *122*(3), 375–391. <https://doi.org/10.1007/s00702-014-1258-3>
- Vanthornhout J., Decruy L., Wouters J., Simon J. Z., & Francart T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *Journal of the Association for Research in Otolaryngology*, *19*(2), 181–191. <https://doi.org/10.1007/s10162-018-0654-z>
- Varao-Sousa T. L., Smilek D., & Kingstone A. (2018). In the lab and in the wild: how distraction and mind wandering affect attention and memory. *Cognitive Research: Principles and Implications*, *3*(1). <https://doi.org/10.1186/s41235-018-0137-0>
- Varma S., & Simon R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(91), 1–8. <https://doi.org/10.1186/1471-2105-7-91>
- Wendt D., Hietkamp R. K., & Lunner T. (2017). Impact of noise and noise reduction on processing effort: A pupillometry study. *Ear and Hearing*, *38*(6), 690–700. <https://doi.org/10.1097/AUD.0000000000000454>
- Winneke A. H., Schulte M., Vormann M., & Latzel M. (2020). Effect of directional microphone technology in hearing aids on neural correlates of listening and memory effort: an electroencephalographic study. *Trends in Hearing*, *24*, 1–16. <https://doi.org/10.1177/2331216520948410>