Effects of directional hearing aid processing and motivation on EEG responses to continuous noisy speech

TOBIAS NEHER[1,*], BOJANA MIRKOVIC[2,3] AND STEFAN DEBENER[2,3]

[1] *Institute of Clinical Research, University of Southern Denmark, Denmark*

[2] *Department of Psychology, University of Oldenburg, Germany*

[3] *Cluster of Excellence "Hearing4all", Oldenburg, Germany*

Arguably, the next frontier in hearing aid (HA) development are devices that can infer (or learn) the needs of the user via non-invasive physiological measurements such as electroencephalography (EEG) and adjust themselves accordingly. A promising approach to translating EEG signals into HA control signals is the analysis of EEG impulse responses to running speech, as obtained by cross-correlating the audio stimulus with the concurrently recorded EEG signal. Here, we used this method for examining neural correlates of the effects of directional HA processing and listener motivation on speech comprehension in noise. Groups of older participants with normal or impaired hearing listened to audiobook material embedded in realistic cafeteria noise while their EEG was recorded using mobile hardware. A HA simulator was used for (dis)engaging a directional microphone setting and for providing amplification. Motivation was manipulated by offering a monetary reward for good speech comprehension in half of the trials. Motivation influenced the participants' listening performance but not their EEG responses. Directional HA benefit, however, was reflected in both the behavioural and EEG data, thereby illustrating the potential of the tested approach with respect to enabling online HA control.

INTRODUCTION

In clinical practice, hearing aids (HAs) are fitted based on the pure-tone audiogram and feedback from the user (Dillon, 2012). Both types of responses are subjective in nature and therefore prone to bias. Physiological measures of the ability to process speech in noise, on the other hand, could provide an objective basis for both hearing assessment and HA adjustment. Recently, a number of studies investigated the potential of electroencephalography (EEG) measurements with respect to HA adjustments (e.g., Bernarding *et al.*, 2017; Van Eyndhoven *et al.*, 2017). A potential advantage of EEG-controlled HAs would be direct access to the cognitive state of the user, which could enable automatic HA adjustments in response to changes in the acoustic environment or the user's intent. However, a prerequisite for controlling a HA based on ongoing EEG signals is a robust neural marker that indexes the relevant cognitive processes reliably. In challenging situations, listeners modulate their attention based on the physical properties of the auditory scene (e.g., changes in the signal-to-noise ratio; SNR) and their interest in specific parts of the scene

---

*Corresponding author: tneher@health.sdu.dk

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ISSN: 2596-5522.

(e.g., their intent or motivation). Thus, when investigating EEG markers for HA control, the influence of both bottom-up (acoustic) and top-down (listener-driven) factors needs to be considered.

The current study investigated EEG correlates of both types of factors with a view towards enabling online HA control. More specifically, it focused on the effects of SNR improvement as brought about by directional HA processing as well as listener motivation with respect to speech comprehension in noise. Below, a summary of the methods and results is provided. Parts of them are based on Mirkovic *et al.* (2019).

METHODS

Participants

Out of 38 recruited participants, 16 normal-hearing (NH; mean age = 67 yrs, range: 62-75 yrs) and 15 hearing-impaired (HI; mean age = 74 yrs, range: 63-88 yrs) participants completed the entire study. The normal-hearing participants were required to have pure-tone average hearing losses as calculated across 0.5, 1, 2 and 4 kHz (PTA4) of less than 25 dB HL in both ears. The hearing-impaired participants were required to have PTA4s of at least 35 dB HL in both ears. Figure 1 shows the average audiograms of these two groups. The HI participants were all bilateral HA users with at least six months of HA experience.

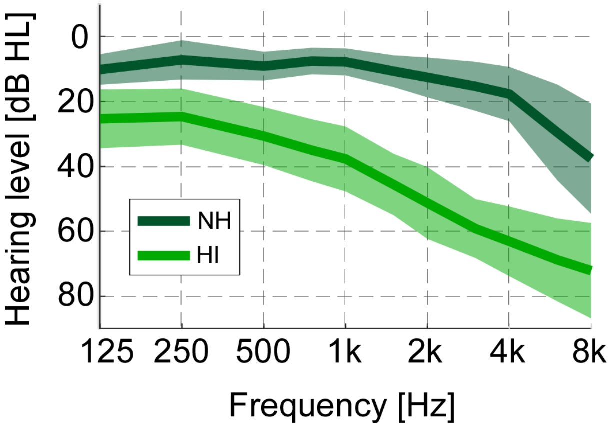

Fig. 1 (colour version online): Mean audiograms of the NH (dark green) and HI (light green) participants. Shaded areas represent ±1 standard deviation (SD).

Experimental paradigm and stimuli

Following the pure-tone audiometry, individual speech reception thresholds (SRTs) were measured using the procedures of Neher *et al.* (2017). The speech material from the Oldenburg sentence test (Wagener *et al.*, 1999) was used as target speech. To create the perception of a spatial auditory scene, the target sentences were convolved with head-related impulse responses recorded in an empty cafeteria with a frontal (0°) source and HA dummies placed on a head-and-torso simulator (Kayser *et al.*, 2009). As background noise, a recording of the fully occupied cafeteria was used. To compensate for the raised hearing thresholds of the HI participants, individual amplification according to the "National Acoustic Laboratories-Revised Profound" prescription rule (Dillon, 2012) was applied to the stimuli using a HA simulator (Grimm *et al.*, 2006). All stimuli were presented binaurally to the participants via insert earphones.

The main task of the participants was to listen to a continuous speech stream masked by the aforementioned cafeteria noise in twelve 10-min long recording sessions. The continuous speech stream consisted of concatenated audiobooks. The cafeteria noise was presented at a nominal level of 65 dB SPL. The target speech was adjusted in level to result in the individual SRTs at the input of the HA simulator. After each recording session, the participants had to answer questions about the contents of the audiobooks. Responses to these questions were used to assess listening performance.

To investigate the effect of motivation, a monetary reward was offered in half of the trials. Below, the two resultant conditions will be referred to as 'motivated' and 'unmotivated'. To investigate the effect of SNR changes, a directional microphone setting was (dis)engaged after each minute of listening. The directional microphone setting corresponded to two (left and right) static forward-facing cardioid microphones. The other setting corresponded to two (left and right) omnidirectional microphones. The directional setting resulted in a speech-weighted SNR improvement of 3.5 dB relative to the omnidirectional setting. Below, these two conditions will be referred to as 'low SNR' and 'high SNR'.

Recording and analysis of EEG signals

While listening to the audiobooks, the participants' EEG signals were recorded using 24 Ag-AgCl electrodes distributed according to the 10-20 system in customised caps. All cap channels were referenced to the FCz channel. The recordings were made with a sampling frequency of 500 Hz using a lightweight, wireless mobile EEG amplifier placed on the back of the EEG cap (see Figure 2).

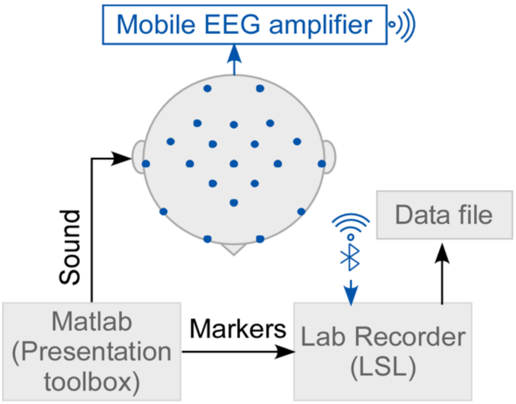

Fig. 2: Illustration of the experimental test setup.

All EEG signal processing was carried out using customized MATLAB scripts as well as the EEGLAB toolbox (Delorme and Makeig, 2004). The raw EEG data were re-referenced offline from the FCz electrode to the algebraic average of the TP9 and TP10 channels to obtain the equivalent of a (commonly used) linked-mastoid reference. Furthermore, the data were band-pass filtered from 0.5 to 45 Hz using a 6th-order Butterworth filter, followed by automatic correction for artefacts such as eye blinks and eye movements using artefact subspace reconstruction (ASR; Mullen *et al*., 2015). Finally, the EEG data were downsampled to 125 Hz.

The EEG impulse responses were estimated based on the analysis pipeline of Petersen *et al*. (2016). First, the Hilbert transform was applied to the speech (or noise) envelope. For each participant, the speech (or noise) was extracted from the signal mixture at the output of the HA simulator and low-pass filtered using a 3rd-order Butterworth filter with a cut-off frequency of 25 Hz. To emphasise speech onsets, the first derivative of the speech envelope was calculated and half-wave rectified. The resultant signal was downsampled to 125 Hz and segmented into 1-min long trials. The analysis pipeline is illustrated in Figure 3. The analysis was performed on the left ear-input signals of each participant.

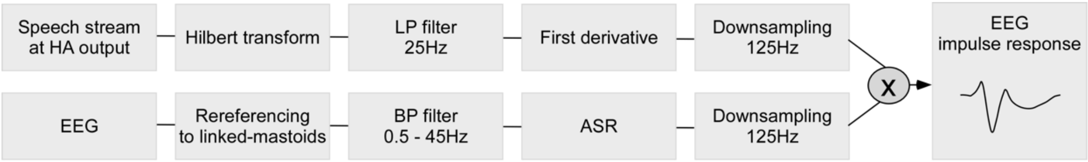

Fig. 3: Illustration of the EEG impulse response analysis pipeline.

Neural oscillations are known to synchronise to the envelope of speech signals – a behaviour that can be modelled by superimposing neural responses to individual speech sounds. The cross-correlation pattern between the resultant EEG envelope and the actual speech envelope reveals the neural impulse response to speech, which is modulated by attention (O'Sullivan *et al*., 2014; Petersen *et al*., 2016). Here, EEG impulse responses were estimated by cross-correlating trials of the extracted speech envelope with corresponding trials of the pre-processed EEG signal, taking into account latencies from -100 to 600 ms. To estimate chance-level synchronisation, EEG trials were also cross-correlated with randomly chosen non-corresponding trials of the speech envelope and used for calculating chance-level EEG impulse responses.

Previous research has found that EEG impulse responses to speech can be physiologically interpreted and that their topography bears similarity to that of auditory evoked potentials (AEPs; Crosse *et al*., 2016). Averages of individual responses were therefore obtained and visualised for the Cz channel where AEPs are traditionally observed. In the resultant grand-average time series, local extremes (or peaks) were identified for each participant, resulting in sets of peak latency and amplitude values that were analysed statistically.

RESULTS

Listening performance

The listening performance of the two participant groups is shown in Figure 4. On average, the NH and HI groups answered, respectively, 68.3% (SD: 12.9%) and 55.3% (SD: 16.7%) of the questions posed correctly. An analysis of variance revealed significant effects of motivation, SNR, group, motivation × SNR and group × motivation (all $p < 0.05$). Closer inspection revealed an influence of motivation in the 'low SNR' ($p < 0.0001$) but not the 'high SNR' ($p > 0.3$) condition. Furthermore, the HI participants performed better in the motivated compared to the unmotivated condition ($p < 0.001$), whereas the NH participants did not ($p > 0.1$).

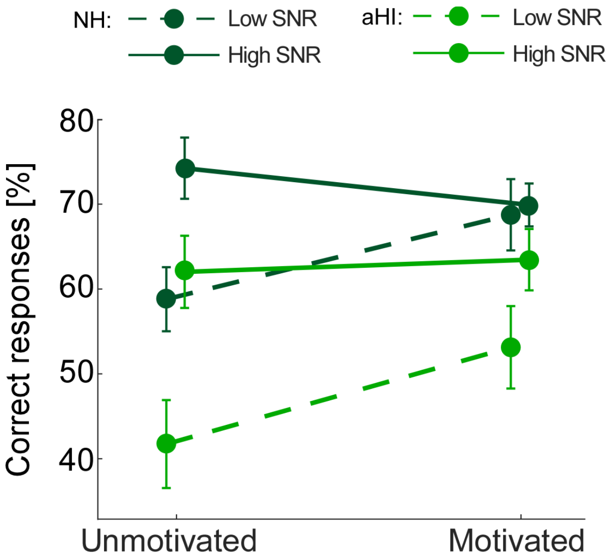

Fig. 4 (colour version online): Mean listening performance of the NH (dark green) and aided HI (light green) groups based on their answers to questions about the contents of the audiobook. Error bars represent ±1 standard error of the mean.

EEG impulse responses

The grand-average EEG impulse responses calculated across all participants are shown in Figure 5. Whereas the chance-level impulse response did not show clear variations in temporal pattern, four peaks were evident in the EEG impulse response for the target speech. The first prominent peak had a latency of 60 ms and was reminiscent of the P100 peak of the traditional AEP response. Responses at this latency are usually characterised as bottom-up-driven. As the focus of this work was on top-down attentional processes, the first peak was not analysed any further. The following analyses were conducted for the other three peaks with latencies of 136 ms ('$\mathrm{N1_{crosscorr}}$'), 240 ms ('$\mathrm{P2_{crosscorr}}$') and 376 ms ('$\mathrm{N2_{crosscorr}}$').

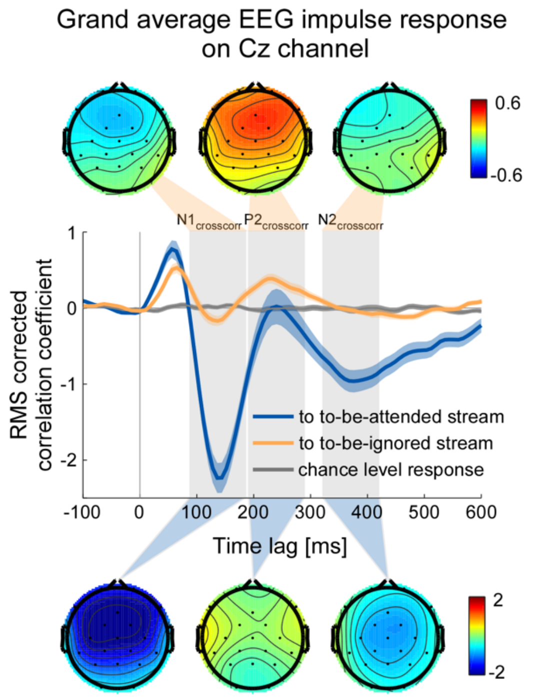

Fig. 5 (colour version online): Grand-average EEG impulse responses for target speech (blue), cafeteria noise (orange) and chance level (grey) across all participants. Shaded areas represent ±1 SD. Topographies corresponding to the grand-average component peaks for the target speech (bottom) and cafeteria noise (top) are also shown.

Analyses of variance carried out on the peak latency and amplitude values revealed significant effects of SNR on N1$_{crosscorr}$ and P2$_{crosscorr}$ latency (both $p < 0.001$) and a significant interaction between group and SNR for the N2$_{crosscorr}$ amplitude ($p < 0.01$). Closer inspection showed faster N1$_{crosscorr}$ and P2$_{crosscorr}$ responses in the 'high SNR' compared to the 'low SNR' condition (see Figure 6). Furthermore, there was a significant influence of SNR on the N2$_{crosscorr}$ amplitude for the HI ($p < 0.01$) but not the NH ($p > 0.1$) participants. Regarding a possible influence of motivation, no effects were observable in any of the N1$_{crosscorr}$, P2$_{crosscorr}$ or N2$_{crosscorr}$ data.

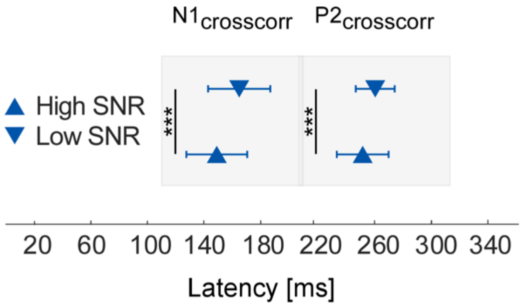

Fig. 6: N1$_{crosscorr}$ and P2$_{crosscorr}$ latencies for the target speech EEG impulse responses when listening with ('high SNR') or without ('low SNR') directional processing. Error bars represent ±1 standard error of the mean.

DISCUSSION

The current study investigated the influence of HA-induced SNR improvement in a realistic auditory scene on both behaviour and neurophysiology. In addition, it manipulated listening motivation using a monetary reward condition. Data were collected using a low-density mobile EEG system and a realistic task, that is, listening to continuous target speech in a spatially complex cafeteria environment. Groups of older NH and HI listeners participated.

As expected, directional HA processing improved listening performance. Furthermore, motivation improved the performance of the HI participants when the directional HA processing was disengaged and listening therefore was demanding. It is currently unclear why the same was not true for the NH listeners who were tested at the same performance level.

The analysis of the EEG impulse responses for the target speech revealed three prominent peaks that are attributable to attentional processes. The topographies and latencies of these peaks corresponded well to those of well-known AEP components (N100 and P200). The latency of these components proved sensitive to the applied SNR changes, with slower responses being evident when the directional HA processing was disengaged. The observed latency effect thus appears to be a physiological measure of the observed behavioural differences.

Regarding the motivational manipulation, no physiological correlates were observable in the EEG data. This was despite the fact that the monetary reward condition improved listening performance, particularly at the low SNR and for the HI group. Thus, further research is needed to investigate top-down influences on EEG responses to continuous speech, perhaps using a different experimental paradigm with a more effective motivational manipulation.

ACKNOWLEDGEMENTS

This research was funded by the DFG Cluster of Excellence EXC 1077/1 “Hearing4all”. The authors thank Julia Schmidt and Manuela Jaeger (University of Oldenburg) for their help with the data collection and data analyses.

REFERENCES

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., and Corona-Strauss, F. I. (**2017**). “Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort,” Cogn. Neurodyn., **11**, 203-215, doi: 10.1007/s11571-017-9425-5.

Crosse, M. J., Di Liberto, G. M., Bednar, A., and Lalor, E. C. (**2016**). “The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli,” Front. Human Neurosci., **10**, 604, doi: 10.3389/fnhum.2016.00604.

Delorme, A., and Makeig, S. (**2004**). “EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” J. Neurosci. Methods, **134**, 9-21, doi: 10.1016/j.jneumeth.2003.10.009.

Dillon, H. (**2012**). *Hearing Aids*, 2nd ed., Thieme.

Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (**2006**). “The master hearing aid: A PC-based platform for algorithm development and evaluation,” Acta Acust. Acust., **92**, 618-628.

Kayser, H., Ewert, S. D., Anemüller, J., Rohdenburg, T., Hohmann. V., and Kollmeier, B. (**2009**). “Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses,” EURASIP J. Advances Sig. Proc., 298605, doi: 10.1155/2009/298605.

Mirkovic, B., Debener, S., Schmidt, J., Jaeger, M., and Neher, T. (**2019**). “Effects of directional sound processing and listener’s motivation on EEG responses to continuous noisy speech: Do normal-hearing and aided hearing-impaired listeners differ?,” Hear. Res., **377**, 260-270, doi: 10.1016/j.heares.2019.04.005.

Mullen, T. R., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.P., and Cauwenberghs, G. (**2015**). “Real-time neuroimaging and cognitive monitoring using wearable dry EEG,” IEEE Trans. Biomed. Eng. **62**, 2553-2567, doi: 10.1109/TBME.2015.2481482.

Neher, T., Wagener, K. C., and Latzel, M. (**2017**). “Speech reception with different bilateral directional processing schemes: Influence of binaural hearing, audiometric asymmetry, and acoustic scenario,” Hear. Res., **353**, 36-48. doi: 10.1016/j.heares.2017.07.014.

O’Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J.J., Shinn-Cunningham, B.G., Slaney, M., Shamma, S.A., and Lalor, E.C., (**2014**). “Attentional selection in a cocktail party environment can be decoded from single-trial EEG,” Cerebral Cortex, **25**, 1697-1706, doi: 10.1093/cercor/bht355.

Petersen, E. B., Wöstmann, M., Obleser, J., and Lunner, T. (**2016**). “Neural tracking of attended versus ignored speech is differentially affected by hearing loss,” J. Neurophysiol., **117**, 18-27, doi: 10.1152/jn.00527.2016.

Van Eyndhoven, S., Francart, T., and Bertrand, A. (**2017**). “EEG-informed attended speaker extraction from recorded speech mixtures with application in neuro-steered hearing prostheses,” IEEE Trans. Biomed. Eng., **64**, 1045-1056, doi: 10.1109/TBME.2016.2587382.

Wagener, K. C., Brand, T., and Kollmeier, B. (**1999**). “Development and evaluation of a sentence test for the German language. I-III: Design, optimization and evaluation of the Oldenburg sentence test,” Z. Audiol. (Audiol. Acoustics), **38**, 4-15, 44-56, 86-95.