PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Realistic audiovisual listening environments in the lab: analysis of movement behavior and consequences for hearing aids

Maartje M. E. HENDRIKSE[1]; Gerard LLORACH[1,2]; Giso GRIMM[1]; Volker HOHMANN[1,2]

[1] Medizinische Physik and Cluster of Excellence 'Hearing4all', Universität Oldenburg, Germany

[2] Hörzentrum Oldenburg GmbH, Oldenburg, Germany

ABSTRACT

With increased complexity of hearing device algorithms a strong interaction between motion behavior of the user and hearing device benefit is likely to be found. To be able to assess this interaction experimentally, more realistic evaluation methods are required that mark a transition from conventional lab experiments to the field. Here we describe audiovisual virtual environments (VEs) that were designed to measure realistic movement behavior in relevant everyday situations in the lab. Movement data was collected from young and elderly normal-hearing participants while they were listening to these VEs. It was shown that the collected data in these VEs is reliable. Furthermore, acoustic simulations are done using the measured head movement in combination with the presented VEs to estimate hearing aid performance related to noise suppression for a set of algorithms. The estimated SNR improvement by the algorithms in the VEs under natural head movement is a better predictor of the everyday life benefit of these algorithms than conventional lab experiments can provide. Moreover, the variance in SNR improvement over the head movement data from the different participants is used to assess the influence of head movement behavior on algorithm performance, showing a strong effect on directional algorithms.

Keywords: audiovisual virtual reality, head movement, hearing aid evaluation

1. INTRODUCTION

In everyday life, people naturally move their head when listening and talking to other people. There are indications that head movement can affect the performance of hearing aid (HA) algorithms. To provide a realistic estimate of the everyday benefit of the HA algorithms, lab experiments should therefore be carried out under natural head movement, which is not the case for conventional lab experiments, e.g. [1]. Based on examples found in literature, two possible consequences of head movement are described that might result in a reduced performance. First, Ricketts [2] showed that head orientation affects the speech intelligibility benefit of directional HA algorithms. This could happen because of **misalignment** between the beam of the directional algorithms and the direction of optimal benefit. Secondly, Abdipour et al. [3] and Boyd et al. [4] provide examples of a reduced performance due to head movement for algorithms that use temporal integration to estimate some property of the acoustic scene (sound source direction). When temporal integration takes place during head movement, it might lead to smeared and therefore inaccurate estimates, potentially resulting in a reduced performance of the algorithm. Furthermore, head movement causes dynamic changes to the acoustic scene to which adaptive algorithms are adapting. All adaptive algorithms have a certain update rate at which they adapt their settings. Instantaneous adaptation is not possible due to artefacts, so there might be some time during adaptation in which the performance is not optimal. Smearing of the estimated acoustic scene properties and the time taken to adapt could be problematic if the algorithm is not fast enough or if the acoustic scene is rapidly changing because of head movement, leading to **maladaptation**.

This study aims to investigate the influence of head movement on HA algorithm performance in terms of noise suppression and relate it to effects of misalignment and maladaptation. In order to investigate this, head movement data are needed in conjunction with the presented audio, separated by

Corresponding author: maartje.hendrikse@uni-oldenburg.de

target and noise. Measuring this in the field is impossible, therefore realistic virtual environments (VEs) that are relevant for everyday life are needed that allow reproducible measurements in the lab. It was shown that visual stimuli are needed in order to measure natural head movement, and that animated characters with certain behavioral features can be used for this [5]. Therefore the VEs have to be audiovisual.

In a previous study [6] such VEs were designed based on relevant everyday situations in order to measure typical movement behavior in the laboratory. Relevant situations were selected based on findings of Eckardt et al., Wagener et al. and Wolters et al. [7–9], and with the aim of implementing a large range of different target sources and distractors. The selected VEs were: a living room, a lecture hall, a train station, a cafeteria and a street. The audio and video of these VEs is published along with a detailed description [10]. In all VEs, the task of the participants was to actively listen to a specified target. The *cafeteria* VE was, in addition to the listening only condition, measured with a dual task condition, where the participants had to listen to the target and at the same time perform a hand-eye coordination task (to simulate listening and eating at the same time). While the participants were listening to the VEs, their head movement was recorded, resulting in a dataset of 21 young normal-hearing and 19 elderly normal-hearing participants that is also published [11]. It was shown that the measured data are reproducible and reliable [6]. The acoustic VEs allow full access to the presented target and noise audio stimuli.

Although the head movement data was measured in normal-hearing listeners, who may behave differently from HA users, we will use it in this study to investigate how natural head movement could influence the performance of hearing aid algorithms. With acoustic simulations based on the measured head movement data and presented audio, an estimate can be obtained of the performance of HA algorithms related to noise suppression in retrospect.

2. METHODS

Acoustic simulations were done using the measured horizontal rotational head movement data from 40 normal-hearing participants (21 young, 19 elderly) and acoustic stimuli from six VEs presented in Hendrikse et al. [6] to estimate the algorithm performance related to noise suppression for a set of standard HA algorithms (Figure 1). The acoustic stimuli from the VEs, separated by target and noise, were put into the acoustic model together with the matching head movement traces. This resulted in an estimate of what HA microphones would have recorded if the participants would have worn a HA while making these head movements. These HA microphone recordings were processed with HA algorithms, resulting in a binaural HA output. From the HA microphone recordings, the input SNR was calculated using segmental SNR after Quackenbush et al. [12] with a window length of 200 ms. From the binaural HA output, the output SNR was calculated, also using segmental SNR. The difference between input and output SNR is the algorithm benefit in terms of SNR improvement. A potential effect of head movement on the algorithm performance would be visible in the variance of the algorithm benefit over the head movement traces. This was related to misalignment and maladaptation by comparing the algorithm benefit for different head directions (misalignment) and for different head speeds (maladaptation). The paragraphs below describe the acoustic model and the HA processing in more detail.

2.1 Acoustic Model

The acoustic VEs were created with the software package TASCAR [13,14]. In TASCAR, a virtual listener was implemented in the VEs who performed the same head movements that had been recorded in the participants. The output of TASCAR were the signals that would have been sent to the loudspeakers for the setup described in Hendrikse et al. [6]. Impulse responses had been recorded from the loudspeakers in this setup to a head-and-torso simulator fitted binaurally with three-channel hearing aid dummies as used in Kayser et al. [15]. Convolution of these impulse responses with the loudspeaker signals resulted in recordings of the HA microphones of the virtual listener while performing the defined head movement.

2.2 Hearing Aid Processing

Six representative HA algorithms from different classes were selected: a delay-and-subtract beamformer (labelled ‘delaysub’), an adaptive differential microphone (‘adm’), a static binaural beamformer (‘beam’), an adaptive minimum variance distortionless response beamformer (‘amvdr’), an adaptive differential microphone ('adm'), a binaural noise reduction algorithm based on interaural coherence ('coh') and a single-channel noise reduction algorithm aiming to identify and enhance speech components ('single'). The implementations of these algorithms in the open Master Hearing Aid (openMHA, version 4.8.0) [16] were used, with their default parameter settings. For processing of the HA microphone recordings with the algorithms the Hagerman method [17] was used, which allowed access to the processed target and noise separately.

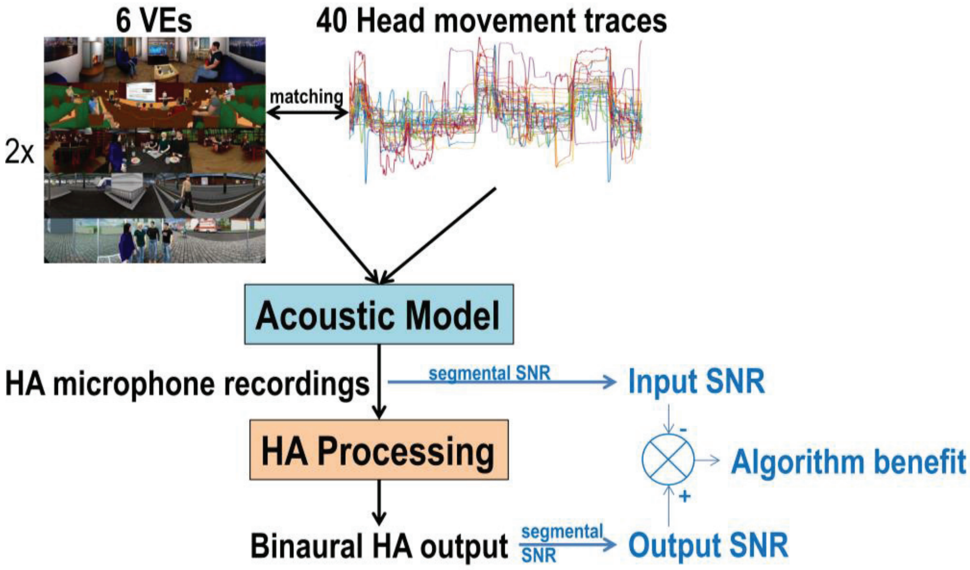

Figure 1: Schematic overview of the applied method. The acoustic stimuli and measured head movement traces from 40 participants for the matching VEs were put into an acoustic model, which provided the HA microphone recordings for a virtual listener performing the head movement. These HA microphone recordings were processed with HA algorithms, resulting in a binaural HA output. The difference in SNR between the HA microphone recordings and binaural HA output is the algorithm benefit.

3. RESULTS

3.1 OVERALL EFFECT OF HEAD MOVEMENT

The average algorithm benefit over time for all tested algorithms in all VEs is plotted in Figure 2. It can be seen that some directional algorithms perform poorly in some VEs, even decreasing the SNR. Figure 2 also shows the average input SNR over time for the different VEs. As can be seen in the figure, the different head movement traces result in a different average input SNR, because of the head shadow effect. The algorithm benefit of the different algorithms is affected by the head movement, too. The difference in algorithm benefit between the different head movement traces can even be up to 4 dB (e.g., 'amvdr' in the *street* VE), but the non-directional algorithms ('coh' and 'single') are only minimally affected. This is the average over time, so the differences at single time points can be higher. The different input SNRs likely play a role in the effect of head movement, but since the range in algorithm benefit is sometimes bigger than the range in input SNR, there are probably also other factors that play a role. The next paragraphs investigate misalignment and maladaptation as potential factors.

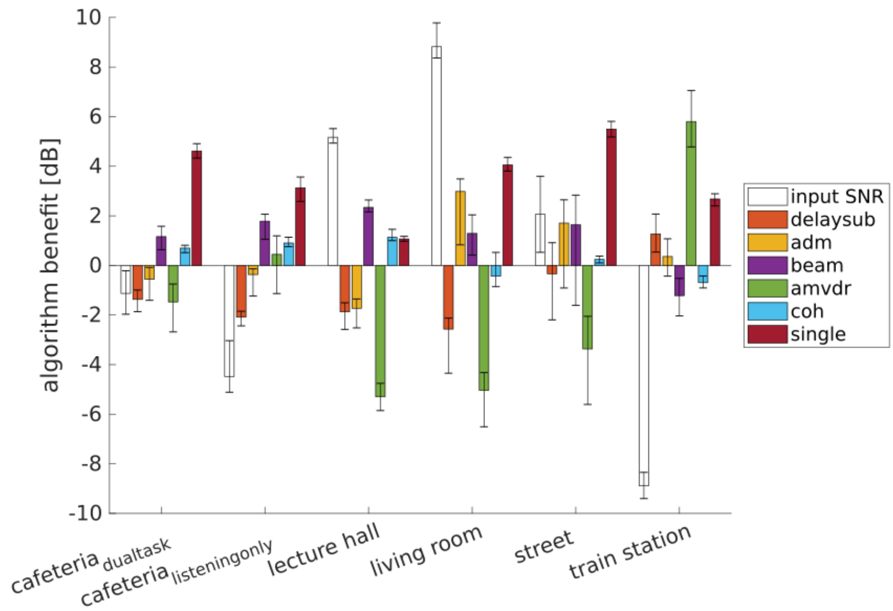

Figure 2: Average input SNR and average algorithm benefit over time for all algorithms in all VEs. The error bar indicates the range over the different movement traces and shows that there is a movement effect.

3.2 MISALIGNMENT

The algorithm benefit per horizontal head direction relative to the target is plotted in Figure 3. In the *train station* VE no target direction was defined, so here the head direction relative to the front is plotted. Misalignment was defined as a decreased performance when a directional algorithm is not pointed in the optimal direction. In case of misalignment problems the algorithm benefit depends on the head direction, as can be seen for all directional algorithms. Especially the 'amvdr' seems affected by misalignment. The 'coh' and 'single' algorithms are non-directional and therefore not affected by misalignment. Their algorithm benefit is fairly constant over different head directions. The values below -70° for the 'single' algorithm in the $cafeteria_{dualtask}$ and *train station* VEs are outliers that are based on very few data points. Misalignment problems are most severe in the *cafeteria* and *street* VEs.

3.3 MALADAPTATION

As explained in the introduction, maladaptation could be a problem if the acoustic scene is changing because of head movement. Maladaptation could therefore cause a decreased algorithm benefit if the head speed is high. The algorithm benefit per horizontal angular head speed is plotted in Figure 4. It can be seen that the algorithm benefit is relatively constant over different head movement speeds for all algorithms. Thus, maladaptation is not a relevant problem.

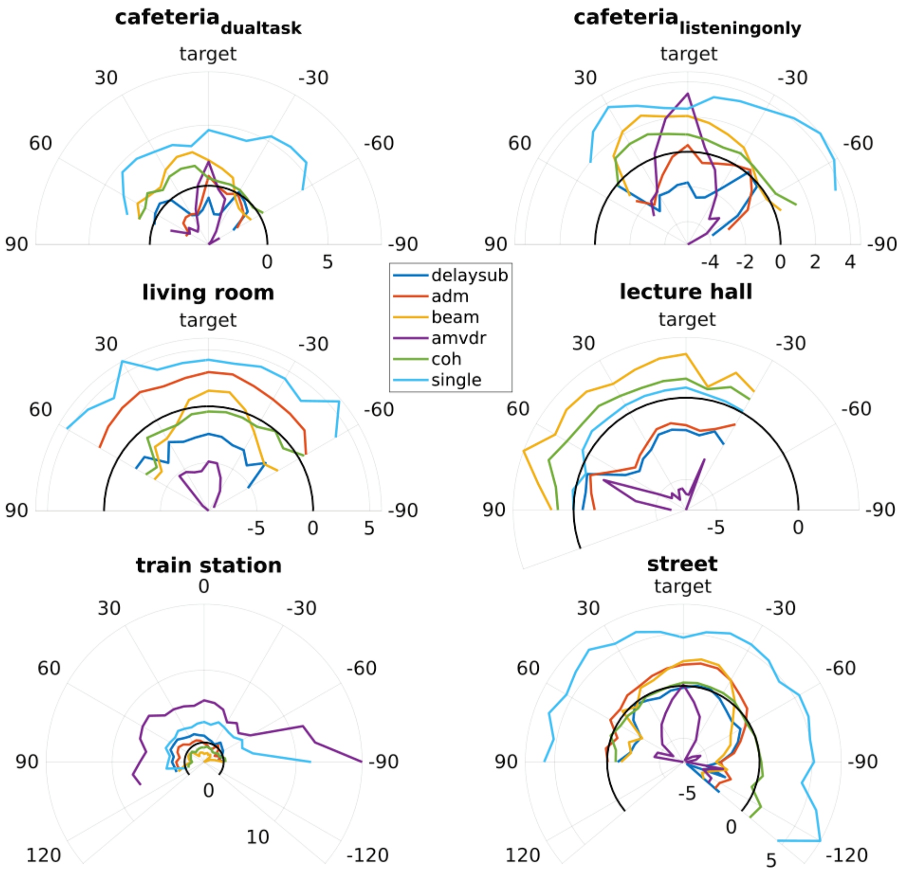

Figure 3: Algorithm benefit per horizontal head direction in degrees relative to the target (*cafeteria$_{dualtask}$*, *cafeteria$_{listeningonly}$*, *lecture hall, living room* and *street*) or relative to the front (*train station*) for all tested algorithms. In case of misalignment problems the algorithm benefit depends on the head direction, as can be seen for all directional algorithms.

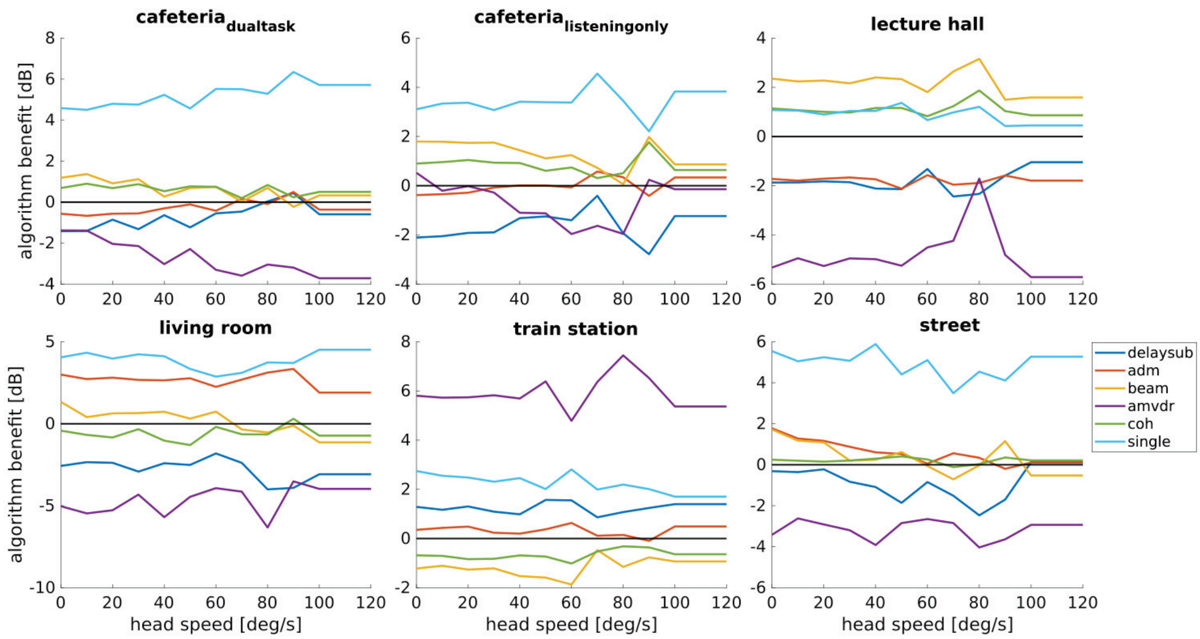

Figure 4: Algorithm benefit per horizontal angular head speed for all tested algorithms in the different VEs. In case of maladaptation problems the algorithm benefit depends on the head speed, which is not the case for any algorithm.

4. DISCUSSION AND CONCLUSIONS

The results show that natural head movement as measured in normal-hearing listeners can severely affect hearing aid algorithm performance. Only directional algorithms are severely affected. Analyses of the causes of differences in performance indicate that both misalignment and differences in the input SNR because of the head shadow effect play a role, and that maladaptation is not a relevant problem. However, it is still unclear how big the relative influences of input SNR and misalignment effect are. Moreover, it would be interesting to see in which situations misalignment problems arise and whether there are differences between the young and elderly participants.

Furthermore, this study shows that directionality of hearing aid algorithms based on the head direction is often not leading to SNR improvement in virtual environments that are relevant for everyday life, based on normal-hearing head movement. In contrary to conventional lab experiments, natural head movements were induced by the visual stimuli and the acoustic complexity of the environments was higher. As shown by Grimm et al. [18], such complex virtual acoustic environments probably lead to a better prediction of everyday benefit. Including head movement could also improve the benefit prediction, because it is shown here that head movement influences the algorithm performance. Although hearing aid users could adapt their head movement behavior in order to increase the benefit, not all of them might do this successfully. Taken together, the hearing aid algorithm performance predicted here is therefore more realistic than the performance predicted in conventional lab experiments. This could help explain the discrepancies found between the directional algorithm benefit measured in conventional lab experiments and the benefit in everyday life [19,20].

Future research could investigate if hearing aid users adapt their head movement and if this leads to an increased algorithm performance. Moreover, if the directional algorithm performance is still poor for adapted head movement behavior, future research should focus on finding ways to make the directionality of the algorithms independent of the head direction. This could be achieved by, e.g., gaze-based attention models [21] or EEG-based attention models [22,23].

ACKNOWLEDGEMENTS

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) –project 352015383 – SFB 1330 B1 and by European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 675324 (ENRICH).

REFERENCES

1. Luts H, Eneman K, Wouters J, Schulte M, Vormann M, Buechler M, et al. Multicenter evaluation of signal enhancement algorithms for hearing aids. J Acoust Soc Am. 2010;127(3):1491.
2. Ricketts T. The impact of head angle on monaural and binaural performance with directional and omnidirectional hearing aids. Vol. 21, Ear and Hearing. 2000. p. 318–28.
3. Abdipour R, Akbari A, Rahmani M, Nasersharif B. Binaural source separation based on spatial cues and maximum likelihood model adaptation. Digit Signal Process. 2015;36:174–83.
4. Boyd AW, Whitmer WM, Brimijoin WO, Akeroyd M a. Improved estimation of direction of arrival of sound sources for hearing aids using gyroscopic information. In: Proceedings of Meetings on Acoustics. Montreal, Canada: Acoustical Society of America; 2013.
5. Hendrikse MME, Llorach G, Grimm G, Hohmann V. Influence of visual cues on head and eye movements during listening tasks in multi-talker audiovisual environments with animated characters. Speech Commun. 2018;101:70–84.
6. Hendrikse MME, Llorach G, Hohmann V, Grimm G. Movement and gaze behavior in virtual audiovisual listening environments resembling everyday life. Submitt to Trends Hear. 2019;
7. Eckardt F, Holube I, Fichtl E, Müller F. Auditory Ecology: Charakterisierung typischer Alltagssituationen mit objektiven und subjektiven Größen. 16 Jahrestagung der Dtsch Gesellschaft für Audiol. 2013;1–5.
8. Wagener KC, Hansen M, Ludvigsen C. Recording and classification of the acoustic environment of hearing aid users. J Am Acad Audiol. 2008;19(4):348–70.
9. Wolters F, Smeds K, Schmidt E, Christensen EK, Norup C. Common Sound Scenarios: A Context-Driven Categorization of Everyday Sound Environments for Application in Hearing-Device Research. J Am Acad Audiol. 2016;27(7):527–40.
10. Hendrikse MME, Llorach G, Hohmann V, Grimm G. Virtual audiovisual everyday-life environments for hearing aid research [Internet]. Zenodo. 2019 [cited 2019 Jan 31]. Available from:

https://zenodo.org/record/1434116#.W9F8_PmYSt9
11. Hendrikse MME, Llorach G, Hohmann V, Grimm G. Database of movement behavior and EEG in virtual audiovisual everyday-life environments for hearing aid research [Internet]. Zenodo. 2019 [cited 2019 Jan 31]. Available from: https://zenodo.org/record/1434090
12. Quackenbush SR, Barnwell TP, Clements MA. Objective measures of speech quality. Prentice Hall; 1988.
13. Grimm G, Luberadzka J, Herzke T, Hohmann V. Toolbox for acoustic scene creation and rendering (TASCAR): Render methods and research applications. In: Neumann F, editor. Proceedings of the Linux Audio Conference. Mainz, Germany: Johannes Gutenberg Universität Mainz; 2015.
14. Grimm G, Luberadzka J, Hohmann V. A Toolbox for Rendering Virtual Acoustic Environments in the Context of Audiology. Acta Acust united with Acust. 2019;105(3):566–78.
15. Kayser H, Ewert SD, Anemüller J, Rohdenburg T, Hohmann V, Kollmeier B. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. EURASIP J Adv Signal Process. 2009;2009.
16. Herzke T, Kayser H, Loshaj F, Grimm G, Hohmann V. Open signal processing software platform for hearing aid research (openMHA). Proc Linux Audio Conf. 2017;35–42.
17. Hagerman B, Olofsson Å. A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. Acta Acust united with Acust. 2004;90(2):356–61.
18. Grimm G, Kollmeier B, Hohmann V. Spatial Acoustic Scenarios in Multichannel Loudspeaker Systems for Hearing Aid Evaluation. J Am Acad Audiol. 2016;27(7):557–66.
19. Cord MT, Surr RK, Walden BE, Dyrlund O. Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. J Am Acad Audiol. 2004;15(April 2003):353–64.
20. Bentler R a. Effectiveness of directional microphones and noise reduction schemes in hearing aids: a systematic review of the evidence. J Am Acad Audiol. 2005;16:473–84.
21. Grimm G, Kayser H, Hendrikse M, Hohmann V. A gaze-based attention model for spatially-aware hearing aids. In: Speech Communication; 13 ITG Symposium. VDE Verlag GmbH Berlin, Offenbach; 2018. p. 231–5.
22. Fiedler L, Wöstmann M, Graversen C, Brandmeyer A, Lunner T, Obleser J. Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. J Neural Eng. 2017;14(3).
23. Mirkovic B, Debener S, Jaeger M, De Vos M. Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. J Neural Eng. 2015;12(4):046007.