Investigating Speech Intelligibility While (Un)Knowingly Listening Through Transparent Hearing Aids

Luca Wiederschein[1], Malte Wöstmann[2], Florian Denk[1], Markus Kemper[1], Hendrik Husstedt[1]
[1] *German Institute of Hearing Aids, Lübeck, Germany, Email: l.wiederschein@dhi-online.de*
[2] *Department of Psychology, University of Lübeck, Lübeck, Germany*

Introduction

Hearing aids enable many people with hearing loss to improve their participation in social life. However, for individuals with normal hearing or mild hearing loss, a hearing aid, or listening through a hearable with a transparent setting, often sounds unfamiliar, which can lead to reduced initial acceptance. It has been shown that in noisy environments, even under transparent conditions (i.e., with the same sound pressure level at the eardrum as in the open ear), hearing aids can introduce a side-effect. This effect negatively impacts speech intelligibility and increases listening effort in normal-hearing individuals [1][2]. It may also partly explain the low acceptance of hearing aids among users with mild to moderate hearing loss and is referred to as the 'Hearing Aid Disadvantage'.

Recent studies [1][3] have demonstrated that this disadvantage persists even when speech and noise originate from the same source, indicating that it can not be entirely attributed to changes in the signal-to-noise ratio caused by the microphone location effect and may instead arise from other factors, such as psychological influences.

This study investigates whether the awareness of wearing a hearing aid influences speech intelligibility. Additionally, listening effort was assessed through self-ratings and pupillometry. A setup was used that enabled the simulation of hearing aid use without participants' awareness. This was achieved by presenting binaurally recorded signals from a head and torso simulator (HATS) over headphones in real time.

Material and Methods

Participants

The study was conducted with 24 young adults with normal hearing (max. 20 dB HL between 250–6000 Hz), aged between 19 and 30 years. The average age of the participants was 23.4 years (SD = 2.89). In total, 14 female and 10 male participants participated. Recruitment was carried out either internally or via a public call on the University of Lübeck's student message board.

Study Setup

The experiment comprised six conditions that varied in both environmental context and presentation mode. The environmental context was either a real or an emulated setting. To this end, the experiment was conducted in two identically constructed, sound-insulated measurement rooms, each equipped with sound-absorbing walls and carpeted floors (T30 < 0.1). In one room, participants were seated 70 cm in front of a Tobii Pro Spectrum eye tracker (300 Hz) and an EIZO EV2451 monitor. A Genelec 8351 loudspeaker was positioned directly above the eye tracker, approximately 30 cm behind it, at 0° azimuth. Sennheiser HDA200 headphones were also placed in this room, with both the loudspeaker and headphones connected to an RME Fireface 802 audio interface. In the second room, a HATS (G.R.A.S. KEMAR), was placed at the same position and orientation relative to a second Genelec 8351 loudspeaker, mirroring the participant's setup in the first room. The HATS was connected via AES3 to a second RME Fireface 802 sound card, which in turn was linked to a MiniDSP processor. This processor transmitted the signal to the audio interface in the first room, enabling real-time audio simulation and loudspeaker equalization. The setup is illustrated in Figure 1.

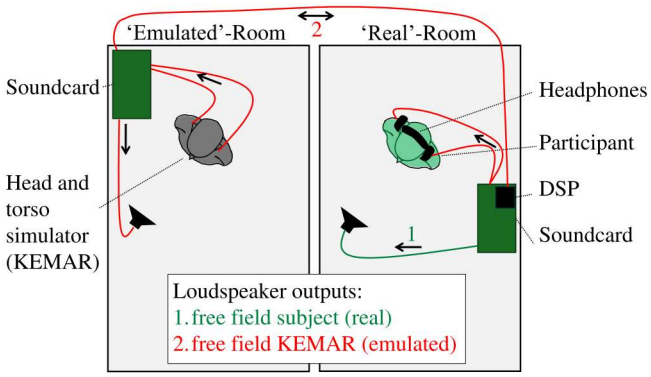

Figure 1: Schematic of the experimental setup with a participant and HATS in rooms of identical construction for real and emulated listening environments.

Participants were measured in three differnt presentation modes. Either using the open ear or a research hearing aid (RHA), the Portable Hearing Laboratory by BatAndCat Sound Labs running the openMHA platform was employed. This platform is well-established in hearing research and has been extensively tested in various experimental settings [5][6][7]. The RHA was fitted according to one of two strategies. First an optimal individually transparent setting based on each participant's real-ear measurement and second a suboptimal generic diffuse field setting based on a previous study [1]. Each of these three configurations, Open Ear, RHA individually transparent, and RHA generic diffuse field, was tested in both distinct environments.

Audio playback was delivered via the loudspeakers in each room, but for the emulated environment, participants listened through the headphones, via the described

audio path from the HATS. For the hearing aid conditions, either the participant (in the real environment) or the HATS (in the emulated environment) wore the RHA. This measurement setup was thoroughly evaluated in a preliminary study [4].

Illumination was standardized to 100 lux at eye level of the participant using a Voltcraft LX-1108 luxmeter to ensure consistent lighting conditions during eye-tracking measurements.

Procedure

At the beginning of the experiment, participants underwent a brief anamnesis, otoscopy, and pure-tone audiometry to ensure normal hearing thresholds. To minimize order effects, the six experimental conditions (three hearing aid configuration in two environments) were presented in a counterbalanced sequence using a 'Latin Square' design.

Speech intelligibility was assessed using the standard adaptive procedure of the Oldenburg Sentence Test (OLSA), a German matrix sentence test typically performed in noise [8]. Initially, participants completed two OLSA training lists to familiarize themselves with the task [8]. The OLSA was used to determine individual values for 50 % speech intelligibility (SRT50) in each condition. All measurements were conducted at a fixed noise level of 70,dB,SPL using the 'Oldenburger Measurement Applications' software.

Listening effort was evaluated by presenting OLSA sentences at two fixed signal-to-noise ratios for each condition. The individually measured SRT50 and an estimated value for 80 % speech intelligibility (SRT80). The SRT80 was derived from an open-ear baseline measurement, assuming a consistent slope of the psychometric function across conditions. Each measurement condition consisted of 40 trials (20 at SRT50 and 20 at SRT80). Each trial began with a baseline window of randomized duration between 1.9 and 3.0 seconds, drawn from a list generated using a right-skewed distribution. Both the pauses and the general measurement procedure was based on a previous study [3] which was based on a best practice approch [9]. This was followed by a randomly selected OLSA sentence lasting between 1.8 and 2.8 seconds, indicated by a loudspeaker symbol on the eye-tracker screen. The sentence was followed by a second randomized pause, generated in the same manner as the baseline window, lasting between 2.6 and 3.7 seconds. In 75 % of the trials, participants were then prompted by a speech bubble on the screen to repeat the sentence within a 4-second response window. This was done to verify that speech intelligibility during the listening effort task remained consistent, did not influence the results, and ensured participants' attention to the task.

Listening effort was evaluated through measurements of pupil dilation relative to prestimuls baseline during the speech test and by administering the NASA Task Load Index questionnaire after each condition [9][10].

Results

Speech Intelligibility Measurements

Figure 2 shows the boxplots of the SNR values for the SRT50, measured using the 'Oldenburg Measurement Applications' across all presentation types in both the real and emulated environments. Lower SNRs indicate better speech intelligibility. Mean SNRs ranged from –6.9 dB (open ear, real) to –5.6 dB (RHA generic diffuse field, emulated), with median values between –6.9 dB and –5.6 dB. The standard deviations ranged from 0.6 dB (RHA generic diffuse field, real) to 1.1 dB (RHA individually transparent, emulated).

A two-way repeated-measures ANOVA revealed significant main effects for both presentation type ($p < 0.001$) and environment ($p = 0.005$), while their interaction was not significant ($p = 0.746$).

Post-hoc (Tukey's HSD) comparisons showed a significant SNR difference between real and emulated environments only for the RHA generic diffuse field condition ($p = 0.019$), with a general trend toward lower SNRs in the real environment for all presentation modes. Differences between presentation modes were significant in both environments, with a 'Hearing Aid Disadvantage' of approximately 0.5 dB for the individually transparent setting and 1 dB for the generic diffuse field setting, regardless of environment.

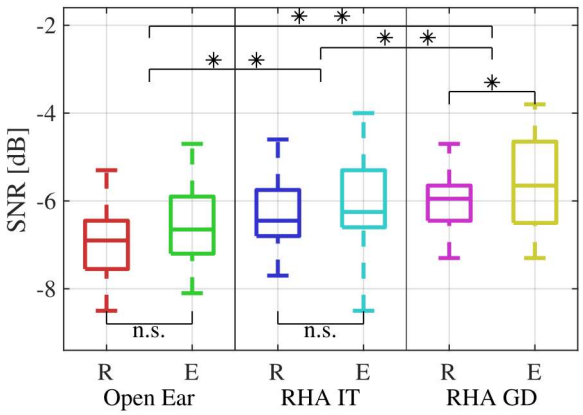

Figure 2: Boxplot of the SNR values for the SRT50 in the real (R) and emulated (E) environments for the presentation modes: Open Ear, RHA individually transparent (RHA IT), and RHA generic diffuse field (RHA GD). Statistically significant differences are indicated by asterisks (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)..

Listening Effort Measurements

Figure 3 shows the average time course of pupil dilation relative to baseline (area increase in percent) for SRT50 trials, aligned to the onset of the OLSA sentence and separated by experimental condition. The maximum pupil dilation (MPD) was calculated as the average within a window of ± 0.5 seconds around the global maximum. The average MPD across participants ranged from 10.3 % (RHA generic diffuse field Real) to 11.3 % (Open Ear Emulated). The time course was similar across all conditions, showing a sharp initial decrease, followed by a peak around 2.6 seconds after sentence onset, and a

second rise during the response window.

No significant effects of environment (Real vs. Emulated) or hearing aid condition (Open Ear, RHA individually transparent, RHA generic diffuse field) were found in terms of MPD. Both paired t-tests and Bayes factors support the null hypothesis, indicating no meaningful difference between real and emulated environments. Similarly, comparisons between open ear and hearing aid configurations showed no consistent pattern and no significant differences. Median changes in pupil size remained nearly identical across all conditions. Finally, no significant difference in MPD was observed between trials with 50% and 80% speech intelligibility ($p = 0.338$).

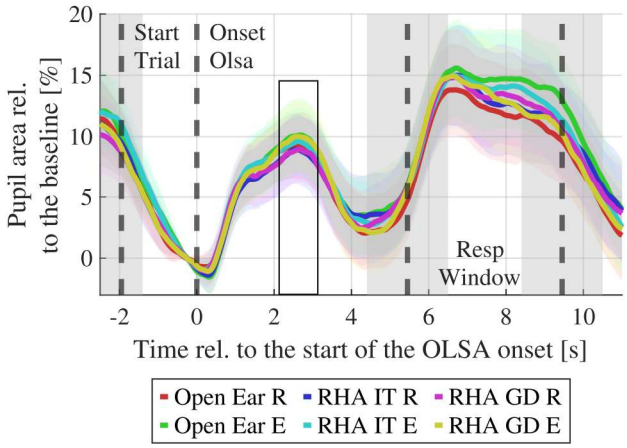

Figure 3: Average progression of pupil area across all participants for trials with a target speech intelligibility of 50 %, shown relative to baseline (%), and time-aligned to the onset of the OLSA trial. Data are separated by condition i.e. the real (R) and emulated (E) environments for the presentation modes: Open Ear, RHA individually transparent (RHA IT), and RHA generic diffuse field (RHA GD). The shaded area represents the standard error of the mean, and the rectangle indicates the time window used to calculate the MPD.

Figure 4 displays a boxplot of responses to the question of the subjective questionnaire, which assessed perceived listening effort. A median 'Hearing Aid Disadvantage' was observed only in the real environment, with response differences of 0.5 scale point for the individually transparent setting and 1 point for the generic diffuse field setting. However, statistical analysis using an ANOVA and post-hoc testing revealed no significant differences between presentation types or environments, except for the comparison between the Open Ear Real and Open Ear Emulated conditions ($p = 0.006$). Thus, no significant 'Hearing Aid Disadvantage' was observed for listening effort using this method.

Discussion

The results regarding speech intelligibility indicate a consistent 'Hearing Aid Disadvantage', in line with previous findings [1][2][3]. In both the real and emulated environments, the SRT50 was consistently higher when using a research hearing aid, with a statistically significant disadvantage of approximately 0.5 dB for the individually transparent settings and nearly 1 dB for the generic diffuse field settings. While overall differences between environments were statistically significant, this effect reached significance at the individual condition level only for the generic diffuse field setting. The results for the open-ear condition align with a preliminary study, supporting the validity of the experimental setup.

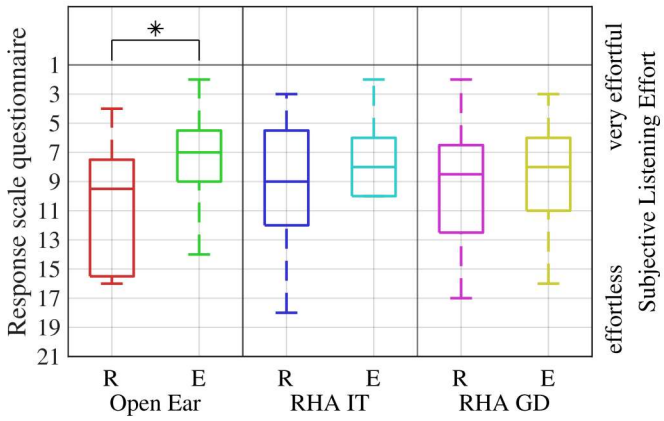

Figure 4: Boxplot providing a detailed visualization of responses to the questionnaire item regarding subjective perceived listening effort. Data are separated by condition i.e. the real (R) and emulated (E) environments for the presentation modes: Open Ear, RHA individually transparent (RHA IT), and RHA generic diffuse field (RHA GD). Statistically significant differences are indicated by asterisks ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$).

No significant effects of environment or presentation mode were found for listening effort in either the pupillometry data or the questionnaire ratings. The only significant difference in the questionnaire data occurred between environments for the open-ear condition. However, certain aspects of the pupillometry data raise concerns about validity, particularly the absence of a difference between the SRT50 and SRT80 conditions, an effect that is well-documented in prior research [11][12]. One possible explanation is a unexpectedly high speech intelligibility during the listening effort task. All conditions, except for the open-ear presentation, exceeded 60 % speech intelligibility, effectively cancelling out the 'Hearing Aid Disadvantage' in this domain. This may be due to learning or habituation effects; however, further investigation is needed to clarify the underlying mechanisms.

The observed difference in perceived listening effort between environments for the open-ear condition may reflect a general discomfort or annoyance caused by wearing equipment, such as headphones or the RHA, on the ear, rather than an effect specific to listening effort.

Conclusion

This study confirms a 'Hearing Aid Disadvantage' in speech intelligibility, consistent with previous studies and observed across both presentation modes. The effect remained stable across both the real and emulated environments. These findings suggest that the 'Hearing Aid Disadvantage' is independent of participants' awareness,

indicating that simply knowing whether one is listening through a hearing aid does not influence speech intelligibility.

In contrast, listening effort results, assessed by pupil dilation and the NASA Task Load Index questionnaire, were inconclusive, with no significant effects observed. The lack of expected differences may be due to methodological factors and warrants further investigation.

References

[1] Denk, F., Wiederschein, L., Kemper, M., Husstedt, H.: (Why) Do Transparent Hearing Devices Impair Speech Perception in Collocated Noise? Trends in Hearing 28 (2024), doi.org/10.1177/23312165241246597

[2] Cubick, J., Buchholz, J. M., Best, V., Lavandier, M., Dau, T.: Listening through hearing aids affects spatial perception and speech intelligibility in normal-hearing listeners. The Journal of the Acoustical Society of America 144 (2018), 2896–2905. doi.org/10.1121/1.5078582

[3] Kemper, M., Husstedt, H., Denk, F., Obleser, J.: Acoustically transparent hearing devices increase parietal alpha power and pupil size in collocated noise at different speech intelligibility levels. Trends in Hearing (2024), in press. doi.org/10.1177/23312165251333225

[4] Denk, F., Bockelmann, A.-S., Tchorz, J., Husstedt, H.: Comparing sound presentation modes for audiological experiments: Loudspeaker, headphones and binaural transmission. 26. Jahrestagung der Deutschen Gesellschaft für Audiologie (DGA 2024), Aalen, Germany. doi.org/10.3205/24dga066

[5] Pavlovic, C., Hohmann, V., Kayser, H., Wong, L., Herzke, T., Prakash, S. R., Hou, Z., Maanen, P.: Open portable platform for hearing aid research. The Journal of the Acoustical Society of America 143 (2018), 1738-1738. doi.org/10.1121/1.5035670

[6] Kayser, H., Herzke, T., Maanen, P., Pavlovic, C., Hohmann, V.: Open Master Hearing Aid (openMHA)—An integrated platform for hearing aid research. The Journal of the Acoustical Society of America 146 (2019), 2879-2879. doi.org/10.1121/1.5136988

[7] Denk, F., Jürgensen, L., Simson, M., Jürgens, T., Husstedt, H.: Performance of research hearing aids: technical evaluation of the portable hearing laboratory. GMS Zeitschrift für Audiologie – Audiological Acoustics 22 (2023). doi.org/10.3205/zaud000020

[8] Wagener, K., Kühnel, V., Kollmeier, B.: Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. Design des Oldenburger Satztests 1 (1999)

[9] Winn, M. B., Wend, D., Koelewijn, T., Kuchinsky, S. E.: Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started. Trends in Hearing 22 (2018). doi.org/10.1177/2331216518800869

[10] Hart, S. G., Staveland, L. E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Advances in Psychology 52 (1988), 139–183. doi.org/10.1016/S0166-4115(08)62386-9

[11] Kuchinsky, S., Vaden, K.: Aging, Hearing Loss, and Listening Effort: Imaging Studies of the Aging Listener. In: The Aging Auditory System (2020), 225–250. Springer. doi.org/10.1007/978-3-030-49367-7_10

[12] Zekveld, A. A., Kramer, S. E.: Cognitive processing load across a wide range of listening conditions: insights from pupillometry. Psychophysiology 51 (2014), 277–284. doi.org/10.1111/psyp.12151