Audio Engineering Society

Conference Paper

Presented at the Conference on
Headphone Technology
2019 August 27 – 29, San Francisco, CA, US

Subjective sound quality evaluation of an acoustically transparent hearing device

Henning Schepker[1,2], Florian Denk[1,2], Birger Kollmeier[1], and Simon Doclo[1]

[1] *Department of Medical Physics and Acoustics and Cluster of Excellence "Hearing4all", University of Oldenburg, Germany*
[2] *These authors contributed equally.*

Correspondence should be addressed to Henning Schepker (henning.schepker@uni-oldenburg.de)

ABSTRACT

In this paper we evaluate the performance of a real-time hearing device prototype that aims at achieving acoustically transparent sound presentation. Acoustic transparency refers to the perceptual equivalence of the sound at the aided ear drum, i.e., with the hearing device inserted and processing on, and the open ear drum, i.e., without the hearing device inserted. The considered hearing device combines a custom earpiece with multiple microphones and signal processing algorithms for robust feedback suppression and sound pressure equalization. We evaluate the perceived overall sound quality of this prototype using dummy head recordings in different acoustic conditions using a multi-stimulus with hidden reference and anchor-like framework with $N = 15$ normal-hearing subjects. Results show that the overall sound quality can be significantly improved for all conditions by using sound pressure equalization, where the processing delay of the device is a crucial limiting factor of the sound quality.

1 Introduction

In the past decades, major improvements have been made in the area of assistive listening devices like hearing aids and consumer headsets. Nevertheless, the acceptance of these devices remains rather limited, with limited sound quality identified as one of the major reasons [1, 2], especially for normal-hearing and mild-to-moderately hearing-impaired subjects. While these people would benefit from advanced signal processing in hearing devices [3], e.g., beamforming, dynamic range compression, and dereverberation, they are usually not willing to accept a reduced sound quality [1]. Therefore, recently the concept of acoustic transparency has become increasingly popular, which aims at creating the acoustic impression of open ear listening while the device is used [4, 5, 6, 7, 8].

Acoustic transparency is achieved when the sound at the aided ear drum, i.e., with the device inserted and processing on, and the open ear drum, i.e., without the device inserted, is perceptually equivalent. Typically, an equalization filter is used to modify the signal picked up by the hearing device such that in superposition with the sound leaking into the (partially) occluded ear canal the desired acoustic characteristics of the open ear are obtained [4, 8]. However, since the output of the hearing device is typically delayed, this superposition may cause comb-filtering effects, possibly degrading the perceived sound quality [9]. For vented hearing devices, it is especially important to take into account the sound leaking into the ear canal [8]. A larger vent typically increases the risk of acoustic feedback and may also reduce the effectiveness of noise reduction algorithms [10, 11, 12]. Hence, for an acoustically transparent hearing device with a larger venting, acoustic feedback suppression is an important component. Since algorithms for equalization and feedback suppression are typically designed independently, combining these algorithms may cause undesired interactions.

The purpose of this paper is twofold. First, we present a prototype real-time acoustically transparent earpiece with integrated single-loudspeaker equalization [13] and multi-microphone acoustic feedback suppression based on an fixed null-steering beamformer [14]. Second, we perform a subjective sound quality evaluation using dummy head recordings and a multi stimulus with hidden reference and anchor (MUSHRA)-like framework [15, 16] with $N = 15$ normal-hearing subjects. We address the following research questions: 1) which equalization algorithm yields the highest perceptual quality compared to the open ear; 2) how is the performance of both the fixed null-steering beamformer and the equalization algorithm affected by different incoming signal directions and reverberation times; 3) how valid is the equalization target used in the equalization filter design; 4) do interactions of the feedback suppression algorithm and the equalization algorithm yield a reduced sound quality.

2 Methods

In Section 2.1 we first present an overview of the considered acoustic hearing device system. We then briefly review the computation of the null-steering beamformer for feedback suppression in Section 2.2 and the equalization for acoustic transparency in Section 2.3. We describe their real-time implementation in Section 2.4. Finally, we describe the experimental setup for the subjective quality evaluation in Section 2.5.

2.1 Hearing Device System Overview

Consider the hearing device system with one loudspeaker and M microphones depicted in Figure 1. This block scheme shows the hearing device processing and all acoustic transfer functions between the source and the eardrum. We assume that all acoustic transfer functions are linear and time-invariant and can be modelled as polynomials in the delay operator q [17].

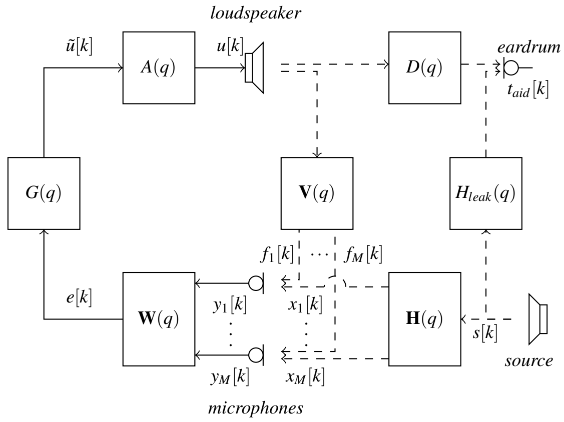

Fig. 1: Single-loudspeaker multi-microphone hearing device system.

The signal $y_m[k]$ in the mth microphone, $m = 1,\ldots,M$, at discrete time k, consists of the incoming signal component $x_m[k]$ and the feedback component $f_m[k]$, i.e.,

$$y_m[k] = x_m[k] + \underbrace{V_m(q)u[k]}_{f_m[k]}, \tag{1}$$

where $V_m(q)$ denotes the acoustic feedback path between the loudspeaker and the mth microphone and $u[k]$ denotes the loudspeaker signal. Furthermore, we assume a single directional incoming signal, i.e.,

$$x_m[k] = H_m(q)s[k], \tag{2}$$

where $H_m(q)$ denotes the acoustic transfer function between the source $s[k]$ and the mth microphone. For convenience we rewrite the microphone signals using vector notation, i.e.,

$$\mathbf{y}[k] = \mathbf{H}(q)s[k] + \underbrace{\mathbf{V}(q)u[k]}_{\mathbf{f}[k]}, \tag{3}$$

where

$$\mathbf{y}[k] = \begin{bmatrix} y_1[k] & \ldots & y_M[k] \end{bmatrix}^T, \tag{4}$$

$$\mathbf{H}(q) = \begin{bmatrix} H_1(q) & \ldots & H_M(q) \end{bmatrix}^T, \tag{5}$$

and $\mathbf{f}[k]$ and $\mathbf{H}(q)$ are defined similarly as $\mathbf{y}[k]$ and $\mathbf{H}(q)$, respectively. The microphone signals are then combined by applying a filter-and-sum beamformer $\mathbf{W}(q)$, i.e.,

$$e[k] = \mathbf{W}^T(q)\mathbf{y}[k]. \tag{6}$$

Although this beamformer is often designed to reduce ambient noise [3], in this paper we will only use the beamformer $\mathbf{W}(q)$ to reduce the feedback component $\mathbf{f}[k]$ in the beamformer output while preserving the incoming signal (cf. Section 2.2). The transfer function from the source to the output of the beamformer is defined as:

$$H_{dev}(q) = \mathbf{W}^T(q)\mathbf{H}(q). \tag{7}$$

The beamformer output signal is then processed by the forward path of the hearing device $G(q)$, yielding

$$\tilde{u}[k] = G(q)e[k]. \tag{8}$$

Aiming to achieve acoustic transparency, an equalization filter $A(q)$ is applied to this signal yielding the loudspeaker signal

$$u[k] = A(q)\tilde{u}[k]. \tag{9}$$

The signal at the aided eardrum, i.e., with the hearing device inserted and processing the signal, is then defined as

$$t_{aid}[k] = D(q)u[k] + \underbrace{H_{leak}(q)s[k]}_{t_{occ}[k]}, \tag{10}$$

where $D(q)$ denotes the acoustic transfer function between the hearing device loudspeaker and the eardrum, $H_{leak}(q)$ denotes the acoustic transfer function between the source and the occluded eardrum, e.g., due to leakage through the vent, and $t_{occ}[k]$ denotes the signal at the occluded eardrum. The desired signal at the eardrum is defined as

$$t_{des}[k] = G(q)H_{open}(q)s[k], \tag{11}$$

where $H_{open}(q)$ denotes the acoustic transfer function between the source and the open eardrum. The goal of an equalization algorithm is then to design the filter $A(q)$ such that $t_{aid}[k]$ is as close as possible to $t_{des}[k]$ (cf. Section 2.3).

2.2 Acoustic Feedback Suppression Algorithm [14]

In order to suppress the acoustic feedback component in the microphones, we use a time-invariant beamformer that steers a null towards the location of the loudspeaker and aims at preserving the incoming signal for a specific direction [14]. In particular, this null-steering beamformer (NS-BF) aims to achieve the following two conditions simultaneously

$$\mathbf{W}^T(q)\mathbf{V}(q) = 0, \tag{12}$$

$$\mathbf{W}^T(q)\mathbf{H}(q) = H_{ref}(q). \tag{13}$$

While the first condition achieves acoustic feedback suppression, the second condition preserves the incoming signal in a reference microphone in the output of the null-steering beamformer. To compute the null-steering beamformer, we will use the robust least-squares-based design procedure proposed in [14]. This procedure requires multiple sets of measurements of the acoustic feedback paths $\mathbf{V}(q)$ as well as a measurement of the acoustic transfer functions $\mathbf{H}(q)$ for a desired incoming direction.

2.3 Sound Pressure Equalization Algorithm [13]

Aiming at achieving acoustic transparency, in [8, 16] an iterative algorithm was used. In this paper we will use the least-squares-based (LS) equalization filter design procedure proposed in [13]. In order for the signal at the aided eardrum $t_{aid}[k]$ in (10) to be equal to the desired signal at the open eardrum $t_{des}[k]$ in (11), the equalization filter $A(q)$ needs to satisfy, using (7) - (9),

$$\underbrace{A(q)D(q)G(q)H_{dev}(q) + H_{leak}(q)}_{\text{aided transfer function}} = G(q)H_{open}(q). \tag{14}$$

To compute the equalization filter, measurements or estimates of the transfer functions $H_{open}(q)$ from the source to the open ear, $H_{leak}(q)$ from the source to the occluded ear, $H_{dev}(q)$ from the source to the microphone(s) and $D(q)$ from the loudspeaker to the eardrum are required.

We will exploit different possibilities for the transfer function $H_{dev}(q)$. To achieve perfect equalization for the considered setup in Figure 1, it should be chosen as $H_{dev}(q) = \mathbf{W}^T(q)\mathbf{H}(q)$. However, when the beamformer is not known a-priori, $H_{dev}(q)$ could be chosen to be the acoustic transfer function between the sound source and a reference microphone of the hearing device, i.e., $H_{dev}(q) = H_{ref}(q)$.

2.4 Real-time Implementation

Both the fixed null-steering beamformer for feedback suppression as well as the equalization filter for acoustic transparency were implemented on the master hearing aid (MHA) [18], which is a software platform for

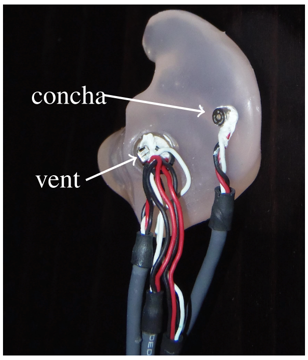

Fig. 2: Custom earpiece used in the hearing device prototype [8]. The microphone in the concha and in the outer side of the vent are indicate. Both loudspeakers are inside the vent, as well as an additional microphone at the inner side of the vent.

real-time signal processing. The MHA was run on an Intel NUC personal computer using an RME Fireface UCX soundcard with a sampling rate of 32 kHz. As earpieces, two custom vented prototypes as described in [8] were used that were inserted in the ear of a dummy head (see Figure 2). The custom earpieces consist of two loudspeakers located in the vent (diameter 4.5 mm, effective diameter due to the transducers ≈1.5 mm) and three microphones, one located at the inner side of the vent in the ear canal, one at the outer side of the vent and one in the concha. Although the earpieces has two loudspeakers, in this study we only use the loudspeaker located at the inner side of the vent. The processing delay of this setup was approximately 6.5 ms, which is in the range of tolerable delays for open fittings [9]. Furthermore, as forward path of the hearing device, a broadband gain $G(q) = 1$ was applied.

In the following we will describe the measurements used for the computation of both the null-steering beamformer $\mathbf{W}(q)$ and the equalization filter $A(q)$. Note that all required acoustic transfer functions could be measured a-priori, e.g., on a dummy head in an anechoic chamber. However, when fitted to human subjects, some of these measurements should be individualized, while others are expected to be less sensitive to individual variations or difficult to measure. Therefore, for some of the acoustic transfer functions we will use estimates, whose influence on the performance will be investigated.

In order to compute the robust null-steering beamformer (cf. Section 2.2), in this study we used two sets of acoustic feedback paths $\mathbf{V}(q)$ per ear of the dummy head, which were measured using sine sweeps [19]. The first set was measured without any objects in the close vicinity of the dummy head and the second set was measured with hands covering the ears. Furthermore, we used a set of acoustic transfer functions $\mathbf{H}(q)$ measured a-priori in an anechoic chamber for a source in front of the dummy head (0°).

Figure 3 shows the directional responses of the left ear for the null-steering beamformer and for the reference microphone at the outer side of the vent, which the null-steering beamformer aims to preserve. In general, the frontal direction is preserved for all frequencies. For different directions the directional responses are very similar for frequencies up to 2000 Hz, while small differences can be observed for higher frequencies. Nevertheless, these results show that the null-steering beamformer does not alter the directional responses of the reference microphone to a large extent.

In order to design the equalization filter (cf. Section 2.3), in this study we performed measurements of the required acoustic transfer functions $D(q)$, $\mathbf{H}_{leak}(q)$, and $H_{dev}(q)$ using sine sweeps [19]. While for $D(q)$ the sine sweeps were played back from the device, in order to measure $\mathbf{H}_{leak}(q)$ and $H_{dev}(q)$ we used Sennheiser HD650 headphones that were put onto the dummy head with the hearing device inserted. Note that a measurement of $D(q)$ in human subjects is difficult and estimation procedures could be used, e.g., based on an electro-acoustic models [20]. However, using electro-acoustic models is beyond the scope of this paper. Furthermore, since the open ear transfer function $H_{open}(q)$ is not available with additional measurement equipment and effort in individual subjects, we used the average diffuse-field equalization function obtained from several subjects from [13] to estimate the open-ear transfer function $H_{open}(q)$ from the acoustic transfer function between the headphones and the concha microphone. Note that the effect of using this estimate will be investigated in the experimental evaluation. For a discussion on the effect of using headphones to measure the required transfer functions, the reader is referred to [8].

Figure 4 shows the aided transfer function of the left hearing device system with feedback suppression and sound pressure equalization when using headphones. As can be observed, the aided transfer function and the estimated open-ear transfer function match well

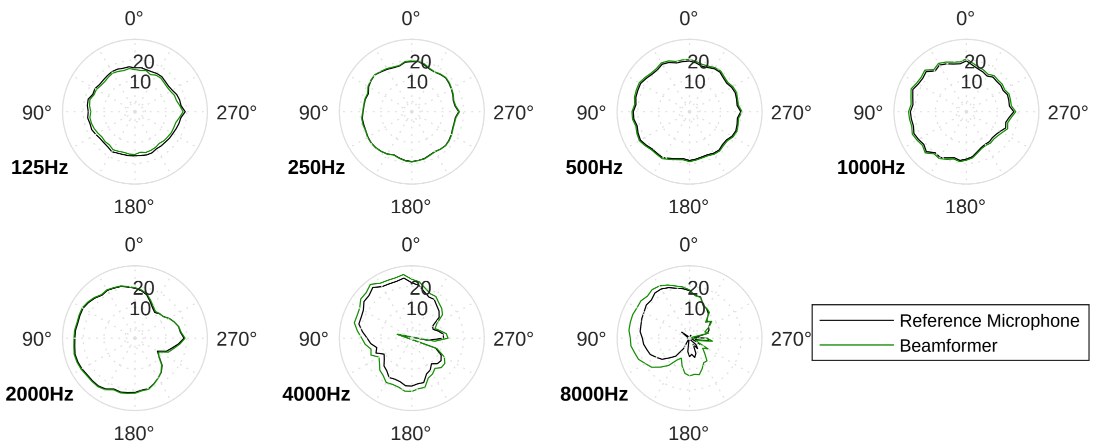

Fig. 3: Directional responses of the left ear for the null-steering beamformer and for the reference microphone for several frequencies.

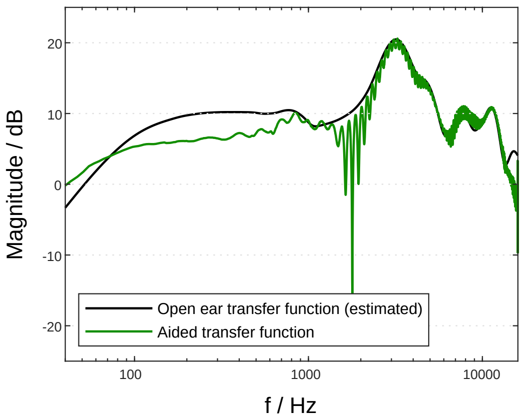

Fig. 4: Open ear transfer function and aided transfer function of the combined system using a null-steering beamformer and the equalization filter in the real-time prototype using headphones.

across the whole frequency range indicating a successful computation of the equalization filter. Nevertheless, comb-filtering effects due to the processing delay are clearly visible that may affect the perceived quality. Note that the comb-filtering effects mainly occur in the frequency range where the leakage component and the output of the hearing device have a similar level.

2.5 Subjective Quality Evaluation

To evaluate the presented acoustically transparent hearing device system, we conducted a formal listening test with $N = 15$ self-reported normal-hearing subjects (none of the authors participated). The task of the subjects was to rate the overall quality of the processed stimuli compared to the open ear reference in a MUSHRA-like framework using a drag-and-drop interface [15]. Note that in contrast to [16], in the present study the reference was explicitly presented to the subjects. Two subjects had to be excluded from the data analysis since they were not able to reliably identify the hidden reference, resulting in a total of 13 subjects (age 28.2±4.0 years). Stimuli were pre-recorded at a sampling rate of 48 kHz using a GRAS 45BB-12 KEMAR Head & Torso with low-noise ear simulators with the hearing device prototype inserted. The dummy head was placed in a lab with variable acoustics (cf. Figure 5), where the reverberation time can be varied using absorber panels mounted at the walls and the ceiling. The recordings were played back to the subjects using MATLAB through an RME ADI-2 Pro FS headphone amplifier and Sennheiser HD650 headphones, which were equalized for a flat magnitude response at the average eardrum.

The goal of the listening test was to assess the impact of different equalization filters and the estimate of the open ear transfer function, the impact of the processing delay of the real-time implementation, as well as potential interactions of the beamformer and the equalization filter. Therefore, the following 8 processing conditions were judged by the subjects (cf. also Table 1), where processing conditions B and C were simulated (without processing delay) and conditions D–G used real-time processing (including a processing delay):

Table 1: Processing conditions used in the experimental evaluation.

Cond.	Feedback Suppression	Equalization Algorithm	$H_{dev}(q)$	Sound Recording	Processing Delay
A	open ear		n/a	$t_{des}[k]$	none
B	none	DF eq. [21]	$H_{ref}(q)$	simulated	none
C	NS-BF [14]	DF eq. [21]	$H_{ref}(q)$	$e[k]$	none
D	NS-BF [14]	Iterative [8]	$\mathbf{W}^T(q)\mathbf{H}(q)$	$t_{aid}[k]$	6.5 ms
E	NS-BF [14]	LS [13]	$\mathbf{W}^T(q)\mathbf{H}(q)$	$t_{aid}[k]$	6.5 ms
F	NS-BF [14]	LS [13]	$H_{ref}(q)$	$t_{aid}[k]$	6.5 ms
G	NS-BF [14]	none	n/a	$t_{aid}[k]$	6.5 ms
H	occluded ear		n/a	$t_{occ}[k]$	none

Fig. 5: Dummy head with inserted hearing device prototypes in a lab with variable acoustics. The green absorbing panels can be flipped to make them highly reflective.

A The open ear reference condition, i.e., without the hearing device inserted to the ear.

B A fully simulated system that uses the measured acoustic transfer functions from the source to the vent microphone $H_{ref}(q)$ and artificially maps it to the open ear using a diffuse field equalization function presented in [21].

C A partly simulated system that uses the output $e[k]$ of the null-steering beamformer for feedback suppression and artificially maps it to the open ear using a diffuse field equalization function presented in [21].

D Using the null-steering beamformer for feedback suppression algorithm and the iterative equalization filter design presented in [8]. Note that the iterative equalization filter design implicitly exploits knowledge about the null-steering beamformer.

E Using the null-steering beamformer for feedback suppression and the LS equalization filter design presented in Section 2.3 computed using $H_{dev}(q) = \mathbf{W}^T(q)\mathbf{H}(q)$.

F Using the null-steering beamformer for feedback suppression presented in Section 2.2 and the LS equalization filter design presented in Section 2.3 computed using $H_{dev}(q) = H_{ref}(q)$.

G Using only the null-steering beamformer for feedback suppression and no equalization filter, i.e., $A(q) = 1$, where the gain of the hearing device was changed compared to conditions D–F to achieve the same broadband level for the aided ear as for the open ear.

H The occluded ear, i.e., with the hearing device prototypes inserted but without processing, providing a low-quality anchor signal.

Using these conditions allows to assess the effect of the diffuse field equalization [21] and possible directional distortions [22] (comparing A and B), the effect of sensor noise (comparing B and C), the effect of the processing delay (comparing C and D, E, and F), and the effect of using equalization (comparing D, E, and F and G). Furthermore, potential interactions of the null-steering beamformer and the equalization filter can be assess (comparing E and F).

The stimuli and acoustic conditions used in the subjective evaluation are shown in Table 2. As stimuli we used two speech signals (male and female) taken from [23] and two music signals (an excerpt from a jazz song [1] and an excerpt from a classical piano recording [2]). The stimuli were played back from three different directions (0°, 90°, 225°) using Genelec 8030 loudspeaker for

[1] J. Redman: Timeless tales for changing times, 1. Summertime
[2] K. Jarret: Bach, Wohltemperiertes Klavier, Book 1, prelude no. 3

Table 2: Overview on acoustic conditions and signals.

Reverberation	Signal Direction	Signals
low $T_{60} \approx 0.35$ s	0°	piano jazz female speech male speech
	90°	piano jazz female speech male speech
	225°	piano jazz female speech male speech
mid $T_{60} \approx 0.45$ s	0°	jazz female speech
	225°	jazz female speech
high $T_{60} \approx 1.4$ s	0°	jazz female speech
	225°	jazz female speech

three different reverberation times: low ($T_{60} \approx 0.35$ s), mid ($T_{60} \approx 0.45$ s), and high ($T_{60} \approx 1.4$ s). The loudspeaker were placed at distance of ≈ 2 m from the dummy head and adjusted in height to be at ear level with the dummy head (approximately 1.6 m). Note that the use of our lab with variable acoustics allowed us to change the reverberation time without changing the physical setup of the loudspeakers and the dummy head.

3 Results

Figure 6 shows the quality ratings (QRs) of the listening test. Individual panels show the results for the different acoustic conditions of Table 2. First, consider the condition with a frontal source (0°) and low reverberation in the top left panel. As can be observed, most subjects were able to identify the hidden reference (processing condition A) and rated the occluded ear with the lowest scores (processing condition H). Furthermore, all signals (piano, jazz, female speech, male speech) were rated similarly for all processing conditions, where generally the piano signal was rated slightly better. Comparing the different processed signals, the fully simulated processing condition B yields the highest QRs that are similar to the open ear, while artificially mapping the output of the null-steering beamformer to the eardrum (processing condition C) yields only a slightly lower QR. Comparing the three real-time processing conditions that include an equalization filter (D–F), it can be observed that all three processing conditions yield similar quality compared to the open ear, with ratings in the range of medium and good QRs. When using no equalization filter (processing condition G), the quality is rated slightly higher compared to the occluded ear (processing condition H). Statistical analyses of the results were conducted using a three-factor analysis of variance (ANOVA) for each of the signal directions with processing condition, reverberation time, and signal as factors using Huynh-Feldt correction for sphericity correction. For the frontal incoming direction the ANOVA showed a significant effect of all three main factors (Proc. condition: $F(2.8, 34.2) = 151.9$, $p < 0.001$; Reverb: $F(0.8, 9.8) = 286.1$, $p < 0.001$; Signal: $F(1.2, 14.7) = 335.6$, $p < 0.001$) as well as their interactions (Proc. condition×Reverb: $F(5.7, 68.4) = 28.1$, $p < 0.001$; Proc. condition×Signal $F(8.5, 102.6) = 46.2$, $p < 0.001$; Reverb×Signal $F(2.4, 29.3) = 219.5$, $p < 0.001$; Proc. condition×Reverb×Signal: $F(17.1, 205.1) = 114.7$, $p < 0.001$). For the 90° direction the ANOVA showed a significant effect of both main factors (Proc. condition: $F(4.1, 49.5) = 142.2$, $p < 0.001$; Signal: $F(1.8, 21.2) = 4.9$, $p < 0.01$) as well as their interaction (Proc. condition×Signal $F(12.4, 148.6) = 2.3$, $p < 0.01$). For the 225° incoming direction the ANOVA showed a significant effect of all three main factors (Proc. condition: $F(2.7, 32.4) = 156.7$, $p < 0.001$; Reverb: $F(0.8, 9.2) = 461.4$, $p < 0.001$; Signal: $F(1.2, 13.9) = 283.5$, $p < 0.001$) as well as their interactions (Proc. condition×Reverb: $F(5.4, 64.7) = 29.9$, $p < 0.001$; Proc. condition×Signal $F(8.1, 97.1) = 51.0$, $p < 0.001$; Reverb×Signal $F(2.3, 27.7) = 178.1$, $p < 0.001$; Proc. condition×Reverb×Signal: $F(16.2, 194.1) = 132.3$, $p < 0.001$). Post-hoc analysis of the three main factors using Bonferroni correction showed for all three signal directions that QRs in low reverberation were significantly higher than in both mid and high reverberation. Furthermore, QRs for the piano signal were significantly higher than for all other signals for the 0° and 225° directions, while for the 90° direction the quality of the jazz signal was rated significantly lower compared to the female speech signal. For all signal directions, QRs were significantly different for all processing conditions, except when comparing processing conditions D, E, and F as well as comparing the processing conditions G and H. Additionally, for the signal directions of 0° and 225°, QRs of processing conditions B and C were not significantly different.

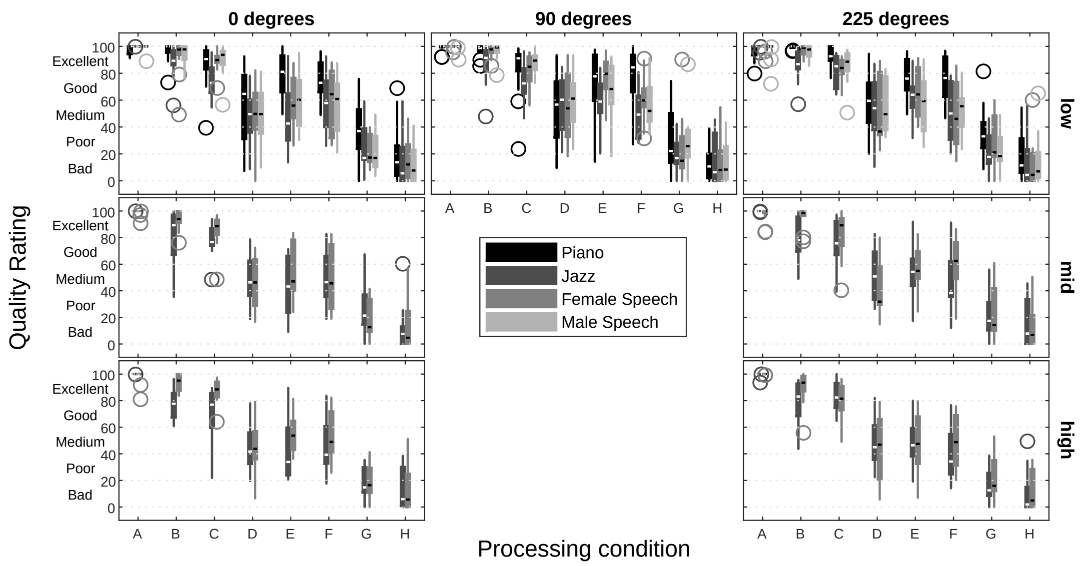

Fig. 6: Results of the formal listening test for different directions (columns) and reverberation times (rows) for the different processing conditions (cf. Table 1) and signals. Lines show the median, boxes show the interquartile ranges, whiskers indicate the last point included within 1.5 times the interquartile range, and circles show outliers.

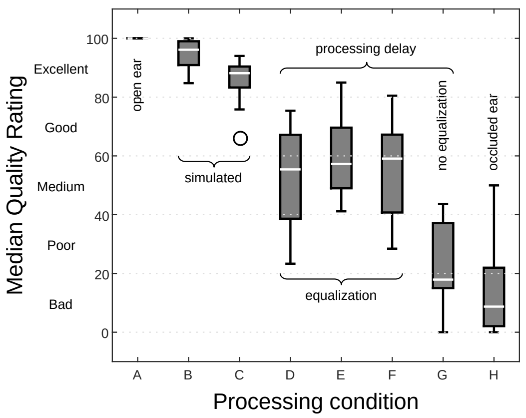

Fig. 7: Median results across all acoustic conditions and signals.

In order to easier visualize the differences between the processing conditions, Figure 7 shows the distribution of the median ratings per subject across all acoustic conditions and signals. Similar trends as in Figure 6 are observed, where a small improvement of the equalization algorithm proposed in [13] (processing condition E) compared to the equalization algorithm proposed in [8] (processing condition D) is observed.

4 Discussion

From the median QRs in Figure 7 it can be observed that, on the one hand, all subjects were able to reliably identify the open ear reference (processing condition A). On the other hand, the occluded ear (processing condition H) was rated worst showing the necessity for sound processing. In the following we will first discuss the results for the real-time processing conditions (D–G). We will then present arguments for the observed significant differences between these processing conditions and the open ear reference condition based on the simulated processing conditions (B, C).

As revealed by the statiscally similar ratings for processing conditions G and H only suppressing the feedback using the null-steering beamformer does not yield a significant improvement compared to the occluded ear. This supports the need for additional processing, i.e., equalization, of the played back signal to achieve a sound quality that is comparable to the open ear, i.e. to achieve acoustically transparent sound presentation. When using an additional equalization filter, a significant improvement in sound quality can be achieved for all considered equalization filters (processing conditions D–F). While there is no significant difference between the different equalization filters, the equalization filter design procedure presented in [13] (processing conditions E and F) is much faster to compute than the iterative procedure presented in [8] (processing condition D). Furthermore, incorporating a-priori knowledge about the null-steering beamformer did not yield an improvement in sound-quality (processing condition C vs D). However, it should be noted that the null-steering beamformer aims at preserving the signal of the reference microphone in its output and thus no large differences were expected.

Even though the proposed hearing device system (processing conditions E–F) achieves a significant improvement in sound quality compared to the occluded ear (processing condition H), the sound quality was still rated lower compared to the open ear (processing condition A). Potential reasons are comb-filtering effects as well as sensor noise. As can be observed by comparing processing conditions A and B, using a precomputed estimate of the open ear transfer function based on the diffuse field equalization [21] is able to achieve almost the same (excellent) quality compared to the open ear. While sensor noise does degrade the quality slightly (processing conditions B vs C), the quality is still perceived as excellent. Comparing processing conditions D–F and G reveals that the processing delay of the real-time system is by far the most important factor that degrades the sound quality. This indicates that when acoustic transparency is desired, the processing delay should be as small as possible to counteract undesired comb-filtering effects. This is especially important in scenarios were the levels of the leakage component and the played back sound are similar, as was the case in the present study where the gain of the hearing device was $G(q) = 1$ (cf. Section 2.4). If the gain of the hearing device is larger, comb-filtering effects are expected to be smaller.

Comparing the different acoustic conditions (reverberation times and signal direction), no large differences can be observed. This indicates that the directional response of the null-steering beamformer (cf. Figure 3) does not largely impact the results. Even though a significant effect of reverberation time was found, the small interaction effect between processing condition and reverberation time indicates that evaluating a limited number of (or even a single) reverberation times is presumably sufficient to investigate difference between hearing devices for acoustically transparent sound presentation.

A similar evaluation was conducted in [16], where equalization using the iterative procedure described in [8] was combined with additional signal enhancement algorithms. In contrast to the present study, however, in [16] the subjects were not provided with an explicit open ear reference and signals were bandlimited to a frequency range of 8 kHz. Results in [16] showed that the quality of transparent sound presentation and the open ear canal were considered similar. Thus it is expected that if subjects were not provided with an explicit reference in the present study, smaller differences between the ratings of the open ear and the presented acoustically transparent hearing device could have been achieved.

5 Summary

In this paper we presented an evaluation of a real-time prototype for acoustically transparent sound presentation. The prototype combines a custom earpiece with multiple integrated microphones with a null-steering beamformer for acoustic feedback suppression and an equalization algorithm taking into account the sound leaking into the ear canal. The results of a formal listening test show that the median QRs of the proposed approach for acoustic transparency are significantly better than the occluded ear (i.e., no processing) and not using sound pressure equalization. Nevertheless, the processing delay of 6.5 ms causing comb-filtering effects, is the main limiting factor for sound quality.

In future work, we aim to investigate the requirements on the processing delay in assistive listening devices with acoustic transparency features as well as using a model-based approach to estimate the sound pressure at the ear drum [20].

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 352015383 – SFB 1330 A4 and C1, and the Cluster of Excellence 1077 "Hearing4all" - Project ID 390895286 - EXC 2177/1.

References

[1] Killion, M. C., “Myths that discourage improvements in hearing aid design,” *Hearing Review*, 11(4), pp. 32–70, 2004.

[2] Sockalingam, R., Beilin, J., and Beck, D. L., “Sound quality considerations of hearing instruments,” *Hearing Review*, 16(3), pp. 22–28, 2009.

[3] Doclo, S., Kellermann, W., Makino, S., and Nordholm, S., "Multichannel signal enhancement algorithms for assisted listening devices," *IEEE Signal Process. Mag.*, 32(2), pp. 18–30, 2015.

[4] Hoffmann, P., Christensen, F., and Hammershøi, D., "Insert earphone calibration for hear-trough options," in *Proc. AES Conf.: Loudspeakers & Headphones, Aalborg, Denmark*, pp. 3–4, 2013.

[5] Rämö, J. and Välimäki, V., "An allpass hear-through headset," in *Proc. Europ. Signal. Process. Conf. (EUSIPCO), Lisbon, Portugal*, pp. 1123–1127, 2014.

[6] Välimäki, V., Franck, A., Rämö, J., Gamper, H., and Savioja, L., "Assisted Listening using a Headset: Enhancing Audio perception in real, augmented and virtual environments," *IEEE Signal Process. Mag.*, 32(2), pp. 92–99, 2015.

[7] Denk, F., Vogl, S., Schepker, H., Kollmeier, B., Blau, M., and Doclo, S., "The Acoustically Transparent Hearing Device: Towards Integration of Individualized Sound Equalization, Electro-Acoustic Modeling and Feedback Cancellation," in *Proc. 1st Int. Workshop Chall. Hearing Assist. Techn. (CHAT), Stockholm, Sweden*, pp. 89–94, 2017.

[8] Denk, F., Hiipakka, M., Kollmeier, B., and Ernst, S. M. A., "An individualised acoustically transparent earpiece for hearing devices," *Int. J. Aud.*, 57(sup3), pp. S62–S70, 2018.

[9] Stone, M. A., Moore, B. C. J., Meisenbacher, K., and Derleth, R. P., "Tolerable Hearing Aid Delays. V. Estimation of Limits for Open Canal Fittings," *Ear and Hearing*, 29(4), pp. 601–617, 2008.

[10] Gatehouse, S., "Limitations on insertion gain with vented earmoulds imposed by oscillatory feedback," *Brit. J. Audiol.*, 23(2), pp. 133–136, 1989.

[11] Dalga, D. and Doclo, S., "Combined feedforward-feedback noise reduction schemes for open-fitting hearing aids," in *Proc. IEEE Workshop on Appl. Signal Process. Audio and Acoust. (WASPAA), New Paltz, New York*, pp. 1–4, 2011.

[12] Winkler, A., Latzel, M., and Holube, I., "Open versus closed hearing-aid fittings: A literature review of both fitting approaches," *Trends in Hearing*, 20, pp. 1–13, 2016.

[13] Denk, F., Schepker, H., Doclo, S., and Kollmeier, B., "Equalization filter design for achieving acoustic transparency in a semi-open fit hearing device," in *Proc. ITG Conf. Speech Comm., Oldenburg, Germany*, pp. 226–230, 2018.

[14] Schepker, H., Nordholm, S., Tran, L. T. T., and Doclo, S., "Null-steering beamformer based feedback cancellation for multi-microphone hearing aids with incoming signal preservation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 27(4), pp. 679–691, 2019.

[15] Völker, C., Bisitz, T., Huber, R., Kollmeier, B., and Ernst, S. M. A., "Modifications of the MUlti stimulus test with Hidden Reference and Anchor (MUSHRA) for use in audiology," *Int. J. Aud.*, 57, pp. S92–S104, 2018.

[16] Denk, F., Kollmeier, B., and Ernst, S. M. A., "High-Fidelity Hearing Instruments: Evaluating Listening Quality of a New Prototype Using a Method for Evaluating Modified Listening (MEML)," in *Proc. AES Int. Conf. Headphone Techn., Aalborg, Denmark*, pp. 1–8, 2016.

[17] Ljung, L., *System Identification: Theory for the User*, Prentice Hall, 1987.

[18] Grimm, G., Herzke, T., Berg, D., and Hohmann, V., "The master hearing aid: a pc-based platform for algorithm development and evaluation," *Acta Acustica united with Acustica*, 92(4), pp. 618–628, 2006.

[19] Müller, S. and Massarani, P., "Transfer-Function Measurement with Sweeps," *J. Audio Eng. Soc.*, 49(6), pp. 443–471, 2001.

[20] Vogl, S. and Blau, M., "Individualized prediction of the sound pressure at the eardrum for an earpiece with integrated receivers and microphones," *J. Acoust. Soc. Am.*, 145(2), pp. 917–930, 2019.

[21] Denk, F., Ernst, S. M. A., Ewert, S. D., and Kollmeier, B., "Adapting Hearing Devices to the Individual Ear Acoustics: Database and Target Response Correction Functions for Various Devices Styles," *Trends in Hearing*, 22, pp. 1–9, 2018.

[22] Denk, F., Ewert, S. D., and Kollmeier, B., "Spectral directional cues captured by hearing device microphones in individual human ears," *J. Acoust. Soc. Am.*, 144(4), pp. 2072–2087, 2018.

[23] Vincent, E., Araki, S., Theis, F., Nolte, G., Bofill, P., Sawada, H., Ozerov, A., Gowreesunker, V., Lutter, D., and Duong, N. Q. K., "The signal separation evaluation campaign (2007–2010): Achievements and remaining challenges," *Signal Processing*, 92(8), pp. 1928–1936, 2012.