Original Article

Adapting Hearing Devices to the Individual Ear Acoustics: Database and Target Response Correction Functions for Various Device Styles

Florian Denk[1] , **Stephan M. A. Ernst**[1,2], **Stephan D. Ewert**[1], and **Birger Kollmeier**[1]

Abstract

To achieve a natural sound quality when listening through hearing devices, the sound pressure at the eardrum should replicate that of the open ear, modified only by an insertion gain if desired. A target approximating this reference condition can be computed by applying an appropriate correction function to the pressure observed at the device microphone. Such Target Response Correction Functions (TRCF) can be defined based on the directionally dependent relative transfer function between the location of the hearing device microphone and the eardrum of the open ear. However, it is unclear how exactly the TRCF should be derived, and how large the benefit of individual, versus generic, correction is. We present measurements of Head-Related Transfer Functions (HRTF) at the eardrum and at 9 microphone locations of a comprehensive set of 5 hearing device styles, including 91 incidence directions, and recorded in 16 subjects and 2 dummy heads. Based on these HRTFs, individualized and generic TRCF were computed for frontal (referred to as free-field) and diffuse-field sound incidence. Spectral deviations between the computed target and listening with the open ear were evaluated using an auditory model and virtual acoustic scenes. Results indicate that a correction for diffuse-field incidence should be preferred over the free field, and individual correction functions result in notably reduced spectral deviations to open-ear listening, as compared with generic correction functions. These outcomes depend substantially on the specific device style. The HRTF database and derived TRCFs are publicly available.

Introduction

Frequency-response characteristics of hearing devices are ideally designed to replicate the individual transfer function to the eardrum of the open ear. In this case, the sound pressure generated at the eardrum of the subject approximates the pressure that would be present at the open eardrum, modified only by an insertion gain if desired. This approach to make hearing devices *acoustically transparent* has been applied to both hearing aids and consumer devices (Denk, Hiipakka, Kollmeier, & Ernst, 2017; Dillon, 2012; Härmä et al., 2004; Hoffmann, Christensen, & Hammershøi, 2013a; Killion, 1979; Rämö & Välimäki, 2012). In the present work, we concentrate solely on the definition of a

suitable target that approaches the open-ear reference, which is independent of the challenge of adjusting the hearing device to create that target at the eardrum of an individual subject. The current article provides an extensive database and analyzes the underlying assumptions and possibilities for deriving correction functions

[1]Medizinische Physik and Cluster of Excellence *Hearing4all*, University of Oldenburg, Oldenburg, Germany
[2]ENT Clinic, University Hospital of Gießen and Marburg, Gießen, Germany

Corresponding author:
Florian Denk, Medizinische Physik and Cluster of Excellence Hearing4all, Universität Oldenburg, Küpkersweg 74, 26129 Oldenburg, Germany.
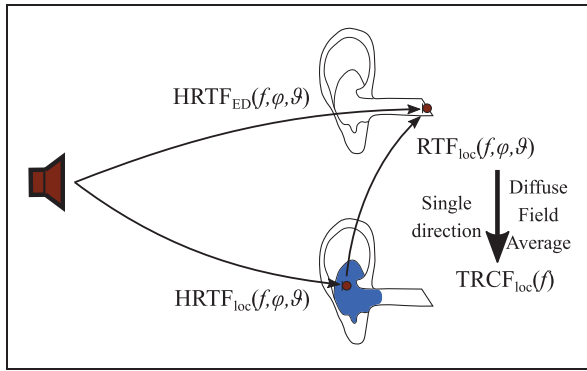Email: florian.denk@uni-oldenburg.de

Figure 1. Illustration of the transfer functions used. To equalize the head-related transfer function (HRTF) to a hearing device microphone (at location loc) to the eardrum of the open ear (ED), the directionally dependent relative transfer function (RTF) between both locations can be applied. φ and ϑ denote the azimuth and elevation angles, respectively. However, in operation, only a directionally independent correction function can be applied, here referred to as target response correction function (TRCF), which can be defined based on the measured directionally resolved RTFs.

that can be employed to define a suitable target and in a comprehensive set of hearing device styles.

The target can either be defined as fixed transfer characteristics of the device as a frequency response, or as a time-varying signal depending on the current input signal. In either case, it is practical to compute the target by transforming the pressure (response) observed at the device's microphone by a frequency-dependent gain. The best possible transformation of the pressure observed at the microphone location of the hearing device to the open eardrum would be the Relative Transfer Function (RTF) between both locations (see Figure 1). For most device styles, the microphone location deviates from an ideal location in the ear canal or at its entrance, which makes the RTF dependent on the direction of incidence (Durin, Carlile, Guillon, Best, & Kalluri, 2014; Hammershøi & Møller, 1996; Hoffmann, Christensen, & Hammershøi, 2013b; Killion & Monsor, 1980). However, without assuming knowledge of the current sound field, only one direction-independent transfer function can be applied. Typical choices are a transformation that is correct for the frontal incidence direction, referred to as free-field correction, or for a random sound incidence, referred to as diffuse-field correction (Bentler & Pavlovic, 1989, 1992; Dillon, 2012; Killion, 1979). In this work, such a transformation will be referred to as the Target Response Correction Function (TRCF). The TRCF is the transformation that corrects the transfer function from the hearing device's microphone to the response to the open eardrum for a given acoustic field and a particular ear. This correction restores acoustic transmission features that are observed at the open eardrum but not at the hearing

device microphone location. The TRCF is a generalization of the CORFIG principle (Coupler Response for Flat Insertion Gain; Killion & Monsor, 1980), which is the TRCF with an additional correction for the hearing device response measured in a 2cc coupler. In contrast to the TRCF, the CORFIG is only applicable in devices with a fully occluding fit, since hear-through sound components are neglected, or it depends greatly on the coupling to the ear. It is well known that the TRCF depends significantly on the hearing device style and microphone location.

Although the TRCF is individual to each ear, in many applications, generic transfer functions must be utilized. The term generic for TRCFs describes nonindividualized corrections that are used for any ear. Whereas in clinical hearing aid fitting the individual TRCF can be measured (implicitly) using probe tube microphone techniques (Dillon, 2012; Mueller, 2001); in many applications, individual measurements at the eardrum are not available. This is the case in self-fit devices, in consumer products, or simply due to cost constraints. Generic transfer functions can be derived from average data of human subjects or dummy head measurements. Bentler and Pavlovic (1989, 1992) have compiled responses from the free and diffuse field to the eardrum and microphone locations of three standard hearing device styles (In-The-Canal, ITC; In-The-Ear, ITE; and Behind-The-Ear, BTE) that were pooled from a large number of separate measurements reported in the literature over several decades (Killion, Berger, & Nuss, 1987; Kuhn, 1979; Kuhn & Burnett, 1977; Madaffari, 1974; Shaw, 1974, 1980; Shaw & Vaillancourt, 1985; Wiener & Ross, 1946). More recently, directionally resolved Head-Related Transfer Function (HRTF) measurements on a dummy head with an ear simulator that included a pair of three-channel BTE hearing aids were presented by Kayser et al. (2009). Durin et al. (2014) provided HRTF measurements on a dummy head and five hearing aid styles with high directional resolution, but excluding transfer functions to the eardrum. However, the existing datasets are limited in terms of device styles and microphone positions in the ear and do not capture differences between individual human ears.

We present HRTF measurements for 91 incidence directions to the eardrum and 9 microphone locations in a comprehensive set of 5 hearing device styles, obtained in both ears of 16 human subjects and 2 commercial dummy heads. The database is publicly available.[1] The data allow extensive analysis of (relative) transfer functions, as well as derivation and evaluation of corresponding TRCFs. Different possible ways to compute the TRCF from the RTFs are evaluated, including individual and average data from human subjects, from dummy head measurements, as well as with free- and diffuse-field corrections. Besides descriptive

analyses of the transfer functions, the expected spectral distortion when listening through TRCF-corrected hearing device HRTFs compared with the open eardrum HRTF was evaluated by means of a psychoacoustic model for linear spectral distortions (Moore & Tan, 2004). Using this approach, we tackled the following research questions relevant to the design and evaluation of hearing devices:

– How can the features of the TRCFs and their dependence on the hearing device style and microphone location be related to known external ear acoustics?
– Is it more appropriate to apply an equalization to the free or to the diffuse field?
– How large is the difference between individual and average correction functions?
– How well can a dummy head-generated TRCF approximate the desired TRCF for the average human listener or a specific individual? Is it beneficial to employ a structural, instead of an arithmetic average?
– What is the putative perceptual relevance of these differences?
– What is the influence of the hearing device style on the TRCF, its difference between individuals and the best directional weighting of RTFs in defining the TRCF?

The article is structured as follows: In the HRTF Measurements section, the measurement routine as well as the hearing device styles and microphone locations are described. In the Analysis Methods section, the further processing of the HRTF data (as published), the RTF extraction, and different possible ways to compute the TRCFs are described. In addition, the method of evaluating the spectral distortion after correction using the different TRCF definitions is described. The Results section shows the measured transfer functions and computed TRCFs, as well as results of the TRCF evaluation. The results are comprehensively interpreted in the Discussion section and the outcomes are summarized in the Conclusions section.

HRTF Measurements

HRTFs for 91 directions were measured in both ears of 16 human subjects (10 male, 6 female, age 27.3 ± 5.1) and 2 dummy heads (*Brüel&Kjær* HATS type 4128C and *G.R.A.S.* KEMAR type 45BM). All experiments were conducted according to the World Medical Association declaration of Helsinki. The subjects were provided with written information material and gave written consent about participation. The dataset contains transfer functions to the eardrum and the microphone locations of a comprehensive set of five hearing device styles. The dataset is publicly available.[1]

Pressure at the Eardrum

The pressure at the eardrum of the open ear was measured by inserting an audiological probe tube microphone until the subject reported contact with the tympanic membrane, and then pulled back by a minimal amount (see Figure 2, panel Eardrum). This procedure (performed by a trained hearing aid acoustician) provided reliable probe tube positioning close to the eardrum, thus minimizing errors due to standing waves in the ear canal in the frequency range of interest (Hellstrom & Axelsson, 1993; Mueller, 2001).

Hearing Device Styles and Microphone Locations

Wide frequency range miniature electret microphones (*Knowles* FG-23329 and GA-38) were used in all hearing devices. The electret microphones were connected to a custom supply and amplifier box providing the operating voltage and 20 dB gain. The number of microphones used was minimized by removable insertion into the different devices whenever possible. All hearing devices with annotated microphone locations are shown in Figure 2. Explanations of the abbreviations for the individual microphone locations are provided in Table 1.

The pressure at the blocked ear canal entrance (ECEbl) was measured with a miniature microphone inserted flush into anthropometric earplugs available in three sizes (Lindau & Brinkmann, 2012), which provide firm and reproducible fit in the ear canal. In a hearing systems context, the blocked ear canal entrance can be regarded as mostly equivalent to small hearing devices fitted into the ear canal of a subject, such as ITC, Completely-In-Canal, or even smaller hearing aids (Bentler & Pavlovic, 1989; Durin et al., 2014).

Another microphone location was on a small insert headphone (InsertHP, *Sennheiser* CX200), as often used in augmented reality audio applications (Härmä et al., 2004; Hoffmann et al., 2013a; Rämö & Välimäki, 2012). A minimal portion of flexible material was attached to the surface, and a miniature microphone inserted flush into a drilled hole. Depending on the subject's ear size, the headphone filled up to half of the cavum conchae. The microphone was placed near the bottom of the concha and pointed toward the rear concha wall (see also Figure 2).

To realize an ITE-type hearing instrument, two microphones were inserted flush into an individual earmold that completely filled the bottom of the concha, one near the ear canal entrance (*Entrance Microphone*, Entr) and one in the rear part of the cavum conchae (*Concha Microphone*). Entrance and concha microphone
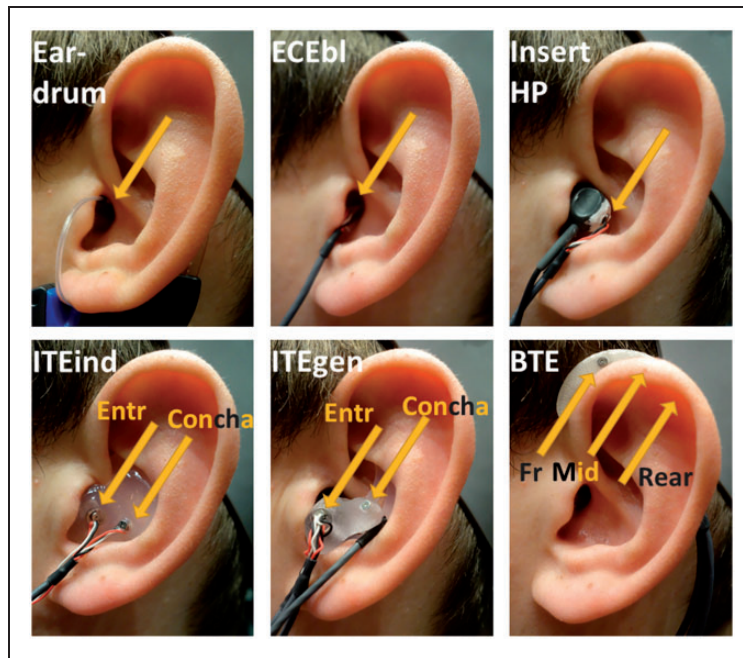
Figure 2. Photograph of all hearing devices and microphone locations in the ear of a subject. Arrows mark the individual microphone locations. Note that two of the sound inlets of the behind-the-ear (BTE) device are behind the pinna. See text for further details, the abbreviations are listed in Table 1.

Table 1. Glossary of Abbreviations.

HRTF	Head-related transfer function
HRIR	Head-related impulse response
DF	Diffuse field
FF	Free field, denoting frontal sound incidence
RTF	Relative transfer function
TRCF	Target response correction function
iDF	Individual, diffuse field
iFF	Individual, free field
mDF	Arithmetic subject average, diffuse field
mFF	Arithmetic subject average, free field
smDF	Structural subject average, diffuse field
dhDF	Arithmetic dummy head average, diffuse field
Hearing device styles/microphone locations:	
ECEbl	Blocked ear canal entrance
InsertHP	Insert headphone with attached external microphone
ITEind, ITEgen	In-the-ear device, *ind*ividual or *generic (nonindividualized form)* earmold
Entr, Concha	*Entrance* and *Concha* microphone
BTE	Behind-the-ear hearing device
fr, mid, rear	*Frontal*, *Middle*, and *Rear* microphone

were approximately 8 to 12 mm apart in a—preferably horizontal—orientation in the individual ears (the distances for the individual subjects are provided with the database). The hardware configuration, referred to as individual ITE device (ITEind), is equal to the outer microphones of the prototype hearing device presented by Denk, Hiipakka et al. (2017).

In a generic ITE device (ITEgen), the microphones were placed in a nonindividualized earplug with a layout comparable to ITEind. A custom adaptor piece was produced that was the same for all subjects and accommodated the transducers that fit into a generic headphone earplug with concha hook (*Bose* StayHear+, one of three sizes selected for each subject). The microphones were 1.1 cm apart and protruded further from the ear than the ITEind earpiece, and the cavum conchae was less uniformly filled. The ITEgen earpiece can also be understood as external microphones contained in a larger insert headphone.

A BTE dummy device with three microphones was produced based on a 3D scan of a commercial hearing aid (the same as used by Kayser et al., 2009). Miniature electret microphones were then placed at the locations of the original sound inlets.

Procedure

The measurements were conducted in an anechoic chamber featuring a 91-channel 3D loudspeaker setup; 48
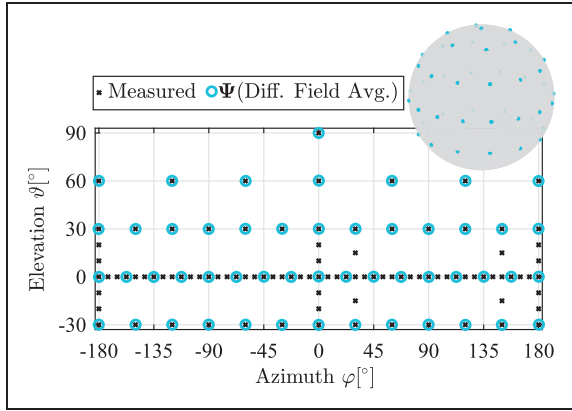
Figure 3. Sound source positions in HRTF measurements in navigational coordinates: All positions measured (black crosses) and those used for diffuse-field averaging Ψ (blue circles). Inset: Sound source positions used for diffuse-field averaging Ψ projected on a sphere.

speakers were located in the horizontal plane leading to a spatial resolution of 7.5°. The rest of the sphere was sampled with a spacing of approximately 30°, but with extra speakers in the median plane and another cone of confusion located at a 30° lateral angle. The spatial sampling grid is shown in Figure 3. Two-way loudspeakers (Genelec 8030b/8020b) were mounted upright between 2.5 and 3 m from the acoustic center. The effects of having separate sound sources for low- and high-frequency reproduction (spacing ca. ± 1.3°) as well as of the varying distance on the HRTF can be neglected (Brungart & Rabinowitz, 1999). In this configuration, it is possible to measure the transfer functions from all different directions simultaneously using overlapping exponential sweeps (Majdak, Balazs, & Laback, 2007). The individual sweeps covered the frequency range between 50 and 20000 Hz with a duration of 4.1 s, leading to a total duration of 36 s for measuring the transfer functions from all loudspeakers. The recordings were made with a sampling rate of 48 kHz. Both the order of directions in each measured HRTF set, as well as the order of device styles for each subject, were randomized. The measurements included one further device style that is included in the public database[1] but not regarded in this work due to close similarities to ITEind. For each hearing device, the measurement was repeated four times without reinsertion of the device. Altogether, the experiment lasted between 60 and 90 min for each subject.

Assessing HRTFs at various points in the ear requires repeated measurements and exchanging the devices between measurements. A particular source of error in this situation is movement of the subject, which would result in HRTF deviations due to an effective shift of the sound incidence direction (Hirahara, Sagara, Toshima, & Otani, 2010). To control this source of inaccuracy, a small headrest in combination with an interactive

positioning method was employed (Denk, Heeren, Ewert, Kollmeier, & Ernst, 2017). The head position was monitored using a headtracker, and necessary corrections to restore a reference position and orientation were displayed to the subject on a graphical interface. This allowed the subject to correct and stabilize their head position with an accuracy of better than 0.5° source shift throughout the experiment, which eliminated the bias caused by variable head orientation. To further reduce the positioning errors, the trial with the best head position was selected for further evaluation, independently for each incident direction. The subjects and dummy heads were initially positioned using a pendulum marking the acoustic room center and a laser distance measurement device. The dummy heads were rotated such that the broadband interaural time difference in the HRTFs to the eardrum for frontal incidence was less than 20 μs (=1 sample @ 48 kHz).

Data Processing

The raw impulse responses contained reflections from equipment, for example, the grating platform the subjects were seated on or the loudspeaker system. These distortions were removed from the data using frequency-dependent truncation (Denk, Kollmeier, & Ewert, 2018). The impulse response was truncated to 4 ms length for frequencies above 1 kHz, but, to avoid truncation errors, not truncated in lower frequency bins, where the reflections did not contain significant energy.

Spectral colorations introduced by the electroacoustic measurement system were compensated by regularized spectral division of the raw frequency responses by the free-field response of every individual loudspeaker (measured with a *Brüel&Kjær* type 4189 microphone), and the individually determined microphone sensitivities. In frequencies exceeding the lower boundary of the measurement (<60 Hz), the responses were extrapolated. To counteract temperature-dependent sensitivity changes of the electret microphones, a broadband gain was applied to each set of HRTFs, to adjust the directionally averaged low-frequency response (average below 150 Hz) to the expected 0 dB. Finally, the resulting impulse responses were truncated to a length of 256 samples, including 16 samples Hann window ramps.

Analysis Methods
Preprocessing and Incidence Directions

For further analysis, HRTFs were calculated from the stored impulse responses with a spectral sampling of 5.9 Hz (8128-point fast fourier transform at 48 kHz sampling rate). Perceptually irrelevant spectral detail was removed by applying complex smoothing of the

spectral power and phase separately, with averaging windows shaped like the responses of 4th-order gammatone filters with 1 ERB bandwidth (Breebart & Kohlrausch, 2001).

A directionally balanced subset Ψ of the measured directions was considered for further evaluation, as indicated in Figure 3. The spacing is 22.5° in the horizontal plane (elevation 0°), 30° in planes with an elevation of $\pm 30°$, 60° in the plane with 60° elevation, and there is one direction at an elevation of 90° (i.e., directly above). This set of $N_\Psi = 47$ incidence directions is approximately equally spaced on the sphere (for elevations $\vartheta \geqslant -30°$, see Figure 3), and an average of transfer functions across these incidence directions can be regarded as an approximation of the corresponding diffuse-field transfer function. Since in the diffuse field, all incidence directions superimpose incoherently, the directional averaging operation was performed on the spectral power values. Phase coefficients were unwrapped and averaged independently.

Relative Transfer Functions

RTF between any given microphone location loc and the eardrum ED in the Free Field (FF) were computed for each subject X and incidence direction separately by complex division of the appropriate HRTFs

$$\text{RTF}_{\text{loc}}^{(\text{FF}, X)}(f, \varphi, \vartheta) = \frac{\text{HRTF}_{\text{ED}}^{(X)}(f, \varphi, \vartheta)}{\text{HRTF}_{\text{loc}}^{(X)}(f, \varphi, \vartheta)}. \qquad (1)$$

As given in the formula, the RTF is, by definition, dependent on the incidence direction; φ and ϑ denote the azimuth and elevation angles, respectively. In the Diffuse Field (DF), the RTF magnitude is given by the quotient of the diffuse field-to-ear transfer functions, approximated by directionally averaged spectral densities

$$\text{RTF}_{\text{loc}}^{(\text{DF}, X)}(f) = \frac{\sqrt{\frac{1}{N_\Psi} \sum_\Psi \left| \text{HRTF}_{\text{ED}}^{(X)}(f, \varphi, \vartheta) \right|^2}}{\sqrt{\frac{1}{N_\Psi} \sum_\Psi \left| \text{HRTF}_{\text{loc}}^{(X)}(f, \varphi, \vartheta) \right|^2}}. \qquad (2)$$

Target Response Correction Functions

The RTF between a hearing device microphone location and the eardrum is the optimal correction function for each incidence direction that must be applied to the microphone signal to obtain the current signal at the eardrum of the open ear. However, despite potential spatial variability of the RTF, only one correction function

can be applied to the input signal. In the following, this correction function is referred to as the TRCF. We define and evaluate different realizations, based on the observed RTFs using individual and averaged data in the human subjects, dummy head data, and different sound-incidence conditions. By averaging across subjects, a transfer function is desired that is correct on average on the dB scale as a rough estimate of the perception. Therefore, TRCF averaging across subjects is conducted in dB values. Note that TRCFs are expressed here by their magnitudes in dB only. The abbreviations for the individual TRCF definitions are summarized in Table 1.

- Individual Free-Field correction, iFF: Utilizing the RTF for a specific microphone location loc for the frontal incident direction in a specific subject's ear X. This is equal to the difference between free-field-to-eardrum and free-field-to-microphone location responses

$$\text{TRCF}_{\text{loc}}^{(\text{iFF}, X)}(f) = 20 \log_{10} \left| \text{RTF}_{\text{loc}}^{(\text{FF}, X)}(f, \varphi = 0, \vartheta = 0) \right|. \qquad (3)$$

- Individual Diffuse-Field correction, iDF: Utilizing the RTF for a specific microphone location loc for diffuse field incidence in a specific subject's ear X. This is equal to the difference between diffuse-field-to-eardrum and diffuse-field-to-microphone location responses

$$\text{TRCF}_{\text{loc}}^{(\text{iDF}, X)}(f) = 20 \log_{10} \left| \text{RTF}_{\text{loc}}^{(\text{DF}, X)}(f) \right|. \qquad (4)$$

- Mean Free-Field correction, mFF: Mean of iFF observed in all N_X human subjects $\mathbf{X}$. This is the free-field correction that is correct for the average of all human subjects

$$\text{TRCF}_{\text{loc}}^{(\text{mFF})}(f) = \frac{1}{N_X} \sum_{\mathbf{X}} \text{TRCF}_{\text{loc}}^{(\text{iFF}, X)}(f). \qquad (5)$$

- Mean Diffuse-Field correction, mDF: Mean of iDF observed in all N_X human subjects $\mathbf{X}$. This is the diffuse-field correction that is correct for the average of all human subjects

$$\text{TRCF}_{\text{loc}}^{(\text{mDF})}(f) = \frac{1}{N_X} \sum_{\mathbf{X}} \text{TRCF}_{\text{loc}}^{(\text{iDF}, X)}(f). \qquad (6)$$

- Structural mean Diffuse-Field correction, smDF: Structural mean of the iDF observed in all subjects. When individual transfer functions are averaged, the resonances tend to be smoothed out, due to variable

peak frequencies across subjects. This can be avoided by finding structural correlates out of which the transfer functions can be computed (ear size etc.), averaging the structural parameters across subjects and computing a structural average (Genuit, 1984; Hammershøi & Møller, 1996; Mehrgardt & Mellert, 1977). As a simple model, two second-order parametric bandpass filters with three parameters each (resonance frequency, resonance gain, and Q-factor, implemented as in Orfanidis [1997]) representing the cavity resonances in the ear canal and the cavum conchae were fitted to each iDF. The resonance frequencies were, however, constrained to sensible boundaries. The structural average transfer function was then computed by averaging the filter parameters across subjects and calculating the corresponding response.

- Dummy-head, Diffuse-Field correction, dhDF: As mDF, but averaged over all dummy head ears, here the KEMAR and Brüel&Kjær HATS.

Evaluation of Spectral Distortion in Corrected HRTFs

Linear spectral distortions compared between listening through the usual HRTF to the eardrum and the hearing device HRTF corrected by a TRCF was evaluated using the model of Moore and Tan (2004). The model calculates the excitation patterns on the basilar membrane and predicts the perceived spectral distortion between a reference and test stimulus (quantified by the metric $D \in [0, 5]$) by evaluating the difference between excitation patterns and the standard deviation along the auditory filters (i.e., spectral ripple). This evaluation was carried out for each ear of the 16 human subjects.

Seven acoustic scenes as introduced by Grimm, Kollmeier, and Hohmann (2016) were used as stimuli. They are spatially rendered speech-in-noise scenes with diverse spectral and spatial energy distribution, and reflect a comprehensive selection of acoustic communication scenarios. The segment between seconds 5 and 10 in each scene was evaluated by the model, with the sound pressure levels adjusted to 75 dB SPL. The same sound files as used by Grimm et al. (2016) were reused, which were rendered for playback on 48 loudspeakers in the horizontal plane. Stimuli referenced to the eardrum were created by convolving the loudspeaker signals with the appropriate Head-Related Impulse Responses (HRIRs). Spectral distortion was evaluated for two cases: For a stimulus with a full audio bandwidth and for a condition in which the acoustic scenes were low-pass filtered at 4 kHz. As a result, only distortions in this low-pass frequency band were captured.

To create the test HRIRs, the smoothed HRTFs to the eardrum and the hearing device microphones were converted back to the time domain and truncated to a length of 400 samples (without loss of information). The TRCFs were then applied by convolving the 200 first samples of their corresponding minimum-phase impulse responses with the HRIRs to the hearing device microphones. Note that during the calculation of the subject-average TRCFs (mFr, mDF, and smDF) used in this evaluation, both ears of a given subject were excluded. The reference HRIRs to the eardrum were created by convolving the hearing device HRIR with the minimum-phase representation of the directionally resolved relative transfer function to the eardrum. This was to circumvent a possible bias in the results through the minimum-phase approximation of the TRCF.

In addition to the perceptive spectral distortion metric D of Moore and Tan (2004), a simpler physical measure based on the spectral difference between the eardrum HRTF- and TRCF-corrected hearing device HRTF was calculated for each condition. The root-mean-square average of the difference spectrum between both HRTFs was computed based on dB magnitudes evaluated in 1/2-ERB spaced auditory filter bands starting at 200 Hz. The RMS was then averaged over all 48 horizontal incidence directions used in the virtual acoustic scenes. The outcome is an easily interpretable spectral difference in dB, which can be compared with the spectral distortion metric for interpretation of the resulting values.

Results

Free and Diffuse Field to Ear Transfer Functions

Figure 4 shows transfer functions from the free field and diffuse field to various points in the ear, including the individual curves for the human subjects, as well as their average. For comparison, comparable data compiled from previous studies (Bentler & Pavlovic, 1989, 1992) are shown.

The current free-field transfer functions are generally in good agreement with literature data. For the eardrum and blocked ear canal entrance (ECEbl), the current data show a systematically higher amplitude in the region of the main resonance (up to 1.5 dB at the eardrum, up to 5 dB at the ECEbl), but the shapes of the curves are very comparable. The free field to ITE and BTE transfer functions from the current study are in excellent agreement with the curves given by Bentler and Pavlovic (1989). At all microphone locations, a slight spectral ripple below 1 kHz is visible in the current data, but not in the curves given by Bentler and Pavlovic (1989). As further assessed in the Discussion section, the ripple most probably originates from a reflection from the legs of the subjects.
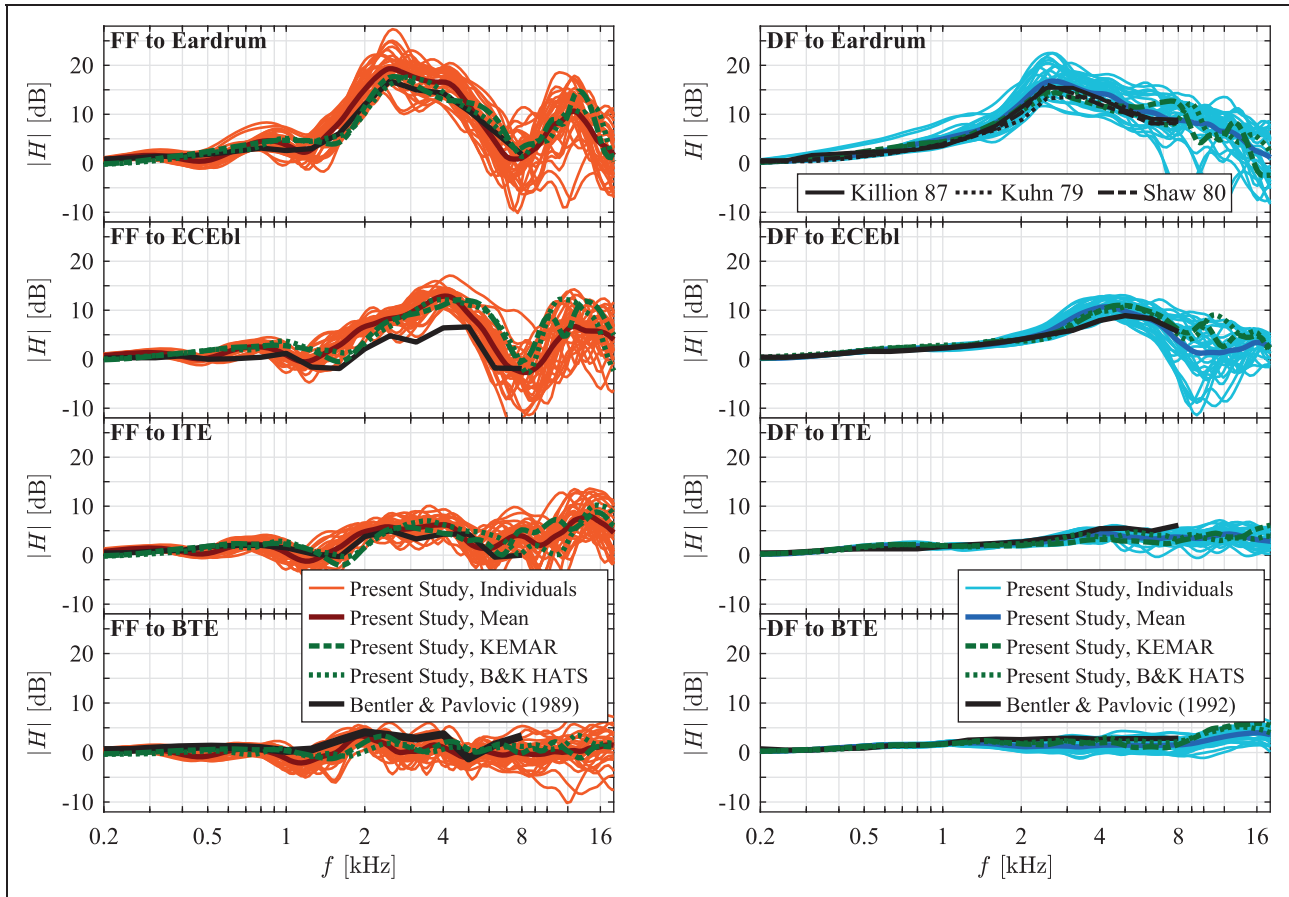
Figure 4. Comparison of free (FF, red curves) and diffuse field (DF, blue curves) to ear transfer functions to literature values (black curves, taken from (Bentler & Pavlovic, 1989, 1992). ECEbl denotes the blocked Ear Canal Entrance, which is compared with the In-The-Canal (ITC) data of Bentler and Pavlovic (1989, 1992), ITE is the Entrance microphone of the individual In-The-Ear device (ITEind_Entr), and for the Behind-The-Ear (BTE) location, the middle microphone (BTE_mid) was selected.

The diffuse-field transfer functions approximations obtained in the current study are in excellent agreement with the diffuse-field curves provided by Bentler and Pavlovic (1992), which were all measured in reverberation chambers. The current diffuse-field-to-eardrum transfer functions are closer to the values reported by Killion et al. (1987) and Shaw (1980) than to those of Kuhn (1979). At the locations ECEbl and ITE, slight systematic differences (< 2 dB) to published values are seen in the higher frequencies.

The difference between subjects is generally larger in the free-field transfer functions. Also, a reduction of the interindividual variations with increasing distance from the eardrum is noted, which is more pronounced in the diffuse-field transfer functions.

Relative Transfer Functions

RTFs for all incidence directions from all microphone locations to the eardrum in the left ears of two exemplary

subjects are depicted in Figure 5. For the eardrum, the HRTF for all incidence directions is shown. VP_E1 is a man with comparatively large ears, whereas VP_N6 is a woman with one of the smallest ears of all subjects included in this study. The connection of the RTF data to acoustic transmission mechanisms in the external ear is addressed in detail in the Discussion section (Corrections Functions Related to External Ear Acoustics section).

At the ECEbl, the difference between RTFs observed at varying incidence directions is small. The RTFs at this location have one common resonance near 2.2 to 3 kHz and further characteristics in the high frequencies.

With increasing distance from the ECEbl, the difference between incidence directions becomes larger—the directional information at all other locations is biased (starting at frequencies >2-4 kHz, depending on the location). Furthermore, the main resonance increases in level and bandwidth, particularly in the frequencies above the peak. A very slight increase of the peak frequency can
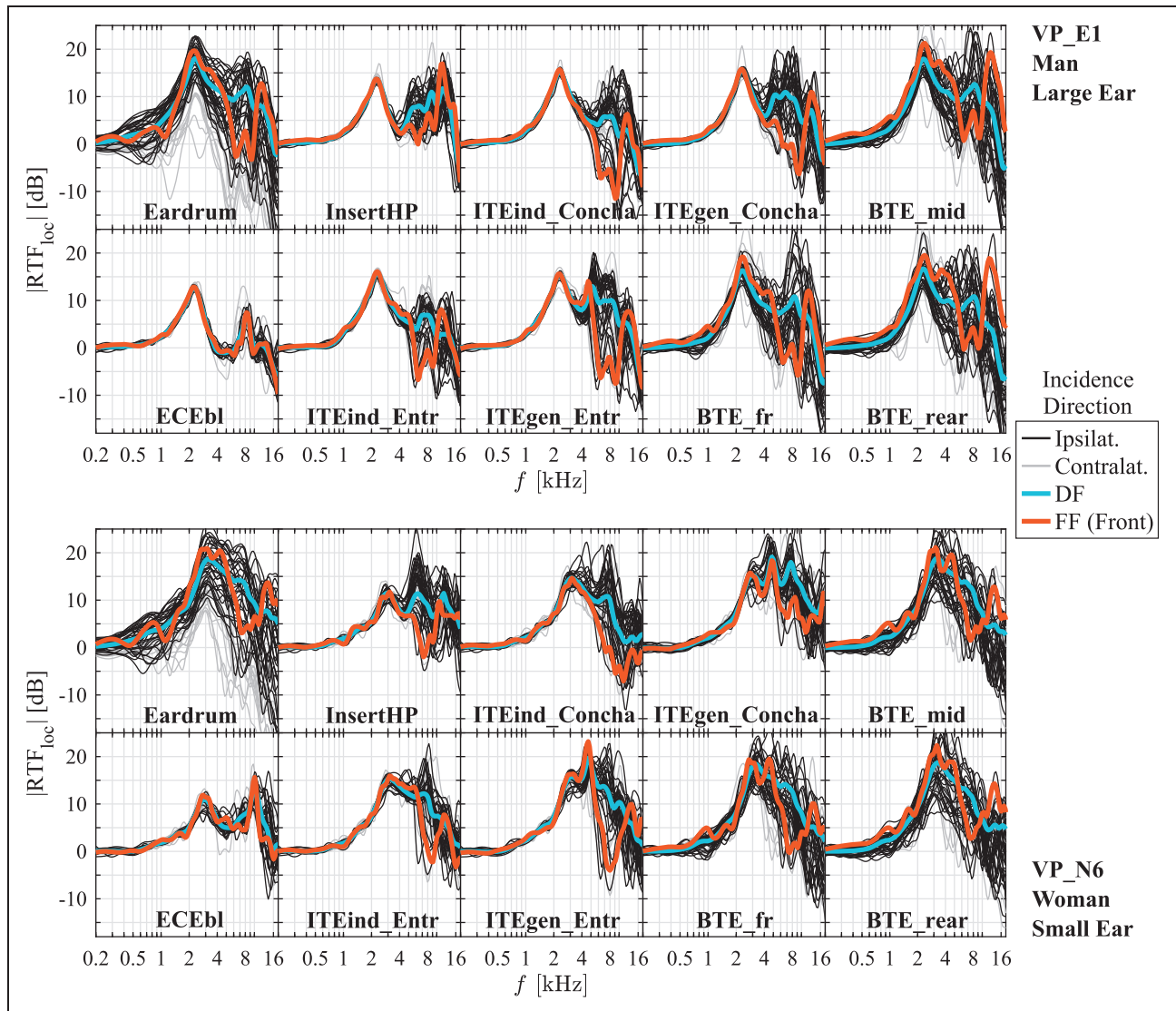
Figure 5. Relative transfer functions (RTF) between all hearing device microphone locations and the eardrum of the open ear for all incidence directions, as well as free-field (FF, i.e., frontal incidence) and diffuse-field (DF) incidence, for two individual subjects. For the eardrum, the corresponding HRTF is shown. VP_E1 is a man with large ears and VP_N6 a woman with small ears.

also be seen. At the BTE microphone locations, the average RTF approaches the average of the eardrum HRTF.

Besides the shape of the main resonance, the RTFs from the different microphone locations also differ in other aspects. Further peaks appear at higher frequencies, being most pronounced in the ECEbl, InsertHP, and ITEgen. Except for the ECEbl, those structures depend greatly on the direction of incidence, and are therefore differently represented in the diffuse-field and the free-field RTF. Between the two curves, the most prominent difference is a dip in the free-field RTF that is not observed in the diffuse-field RTF—and visible in all RTFs except at the ECEbl. Also, a spectral ripple at low frequencies (<1 kHz) is noted for the free-field RTF,

but not in the diffuse field. For both subjects, this ripple is clearly visible in the HRTF to the eardrum and attenuated in the RTF of only some locations, which are not coherent across the subjects. Generally, this feature is more apparent in the RTF data of subject VP_N6.

In the HRTF to the eardrum, a systematic level difference is observed between ipsi- and contralateral incidence directions that is not present in the RTFs. Only at the BTE locations, a slight variation becomes apparent for some directions, particularly toward higher frequencies.

Clear differences between the two subjects are apparent. The most obvious difference between both RTF sets is a shift in the frequency of the first resonance. Also,

relative amplitudes of features at higher frequencies, as well as the spread between different incidence directions, vary notably between the ears shown.

Differences are observed between the RTFs for different microphone locations in the same device. This is particularly true for the free-field RTF (and other individual directions) but also for the diffuse-field average. The difference between the three BTE microphones is rather small, whereas it is largest between the microphones on the ITEgen.

Target Response Correction Functions

Figure 6 shows the TRCF derived from the RTF as defined by Equations (3) to (6) and listed in Table 1. The depiction includes all iDF TRCFs, as well as averages over subjects or dummy heads. Again, prominent variability between subjects is apparent in the iDF curves. The main resonance varies in a range of almost 2 kHz in frequency, as well as about 10 dB in level. At the higher frequencies, differences as large as 20 to 30 dB are observed between the individual iDF-TRCFs. The average Free-Field correction mFF is well in line with the DF curves up to about the first resonance, but clearly deviates from them above ca. 6 kHz. Around 7 kHz, the mFF-TRCF (except at the ECEbl) includes a spectral dip that is not observed in the diffuse-field curves. Only at the BTE microphones, the main resonance is higher in amplitude by a few dB in the mFF than in the mDF correction function. Whereas at most microphone locations, one prominent (broadened) peak is noted in the

mid frequency range, and two separate peaks are observed for both ITEgen microphone locations in almost all TRCF versions.

The structurally averaged correction function (smDF) differs from the arithmetic average between subjects (mDF), showing a main resonance that is slightly higher in level (1 to 4 dB, depending on the microphone location) and a generally smaller amount of spectral detail as a consequence of the two-resonator model used to calculate the structural average. The diffuse-field correction observed in the dummy heads (dhDF) is generally similar to the average of human subjects, but is typically lower in level (2 to 5 dB), especially around the main resonance. This level difference increases in the TRCFs with increasing distance from the eardrum.

In Figure 7, all TRCFs are shown for subject VP_E1 and the ITEind_Entr location. Discrepancies between individual and average curves, as well as between FF and DF correction functions become apparent. First, the frequency of the first resonance is different in the individual (indicated by an i) and generic TRCFs—it is lower in this individual ear than for the average (mFF, mDF, smDF) and dummy head (dhDF) curves. Therefore, in this subject, using a generic response (from a dummy head or a subject average) would introduce a clear error due to the shifted resonance. Second, in the dummy head and human-average curves, the main peak is broader than in the individual data. Compared with this subject, the structural average (smDF) better conserves the shape and bandwidth. Up to about 4 kHz,
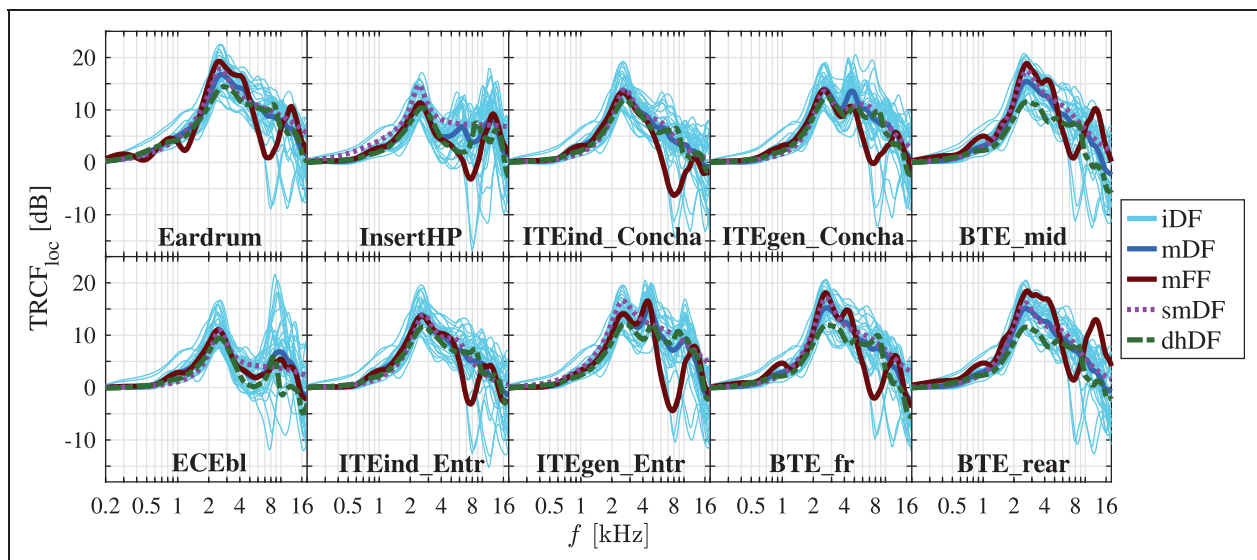

Figure 6. Target response transfer function (TRCF) for each hearing device microphone location, showing the data of individual subjects (only diffuse field, iDF), together with the generic curves obtained from averages of subject and dummy head data. For the eardrum, free or diffuse-field transfer function is shown. See Table 1 for an explanation of the abbreviations.
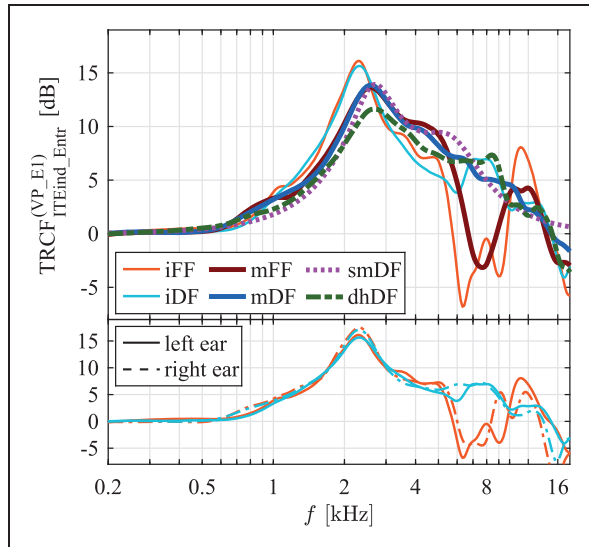
Figure 7. Top Panel: Individual and generic target response transfer functions (TRCF) for one sample ear (left ear of subject VP_E1, as in top panel of Figure 5). Bottom Panel: Individual TRCFs for both ears of the same subject, shown for free- (FF) and diffuse-field (DF) incidence. The line colors in the lower panel are specified in the legend of the top panel. See Table 1 for an explanation of the abbreviations.

Figure 8. Structural parameters of the TRCF main resonance for each hearing device microphone location, distribution over individual diffuse field (iDF) TRCF and results obtained with the generic curves. For the eardrum, the feature distribution of the diffuse field transfer function as shown in Figure 4 is depicted. See Table 1 for an explanation of the abbreviations.

individual free- and diffuse-field TRCFs are virtually identical. At higher frequencies, more pronounced features appear in the free-field TRCF, most prominently a dip around 7 kHz. This and other directionally dependent cues are averaged out in the iDF-TRCF, which at the high frequencies (> 4 kHz) is rather similar to the averaged responses.

The individual TRCFs (iFF and iDF) are very similar in both ears, as shown in the lower panel of Figure 7. An offset at the main resonance frequency of 2 dB between the sides is observed in both ears, and very similarly in the iFF and iDF correction function. At the higher frequencies, the correction functions between ears are very similar in structure, but deviate from each other in the fine details. The difference between ears is larger for the iFF TRCF. Similar observations were made for the other human subjects.

Figure 8 displays the distribution of characteristic parameters of the main resonance in the TRCFs. Distributions of the resonance frequency, the gain, and the bandwidth (±3dB around the maximum) are shown for the iDF TRCF in comparison to the appropriate parameters of the average transfer functions. The increase in gain, bandwidth, and frequency with increasing distance to the eardrum that was already noted in Figure 6 is now clearly visible in the distribution of individual data and to a lesser extent in the generic curves as well. High interindividual variances (spread of iDF data), particularly for resonance frequency and gain,
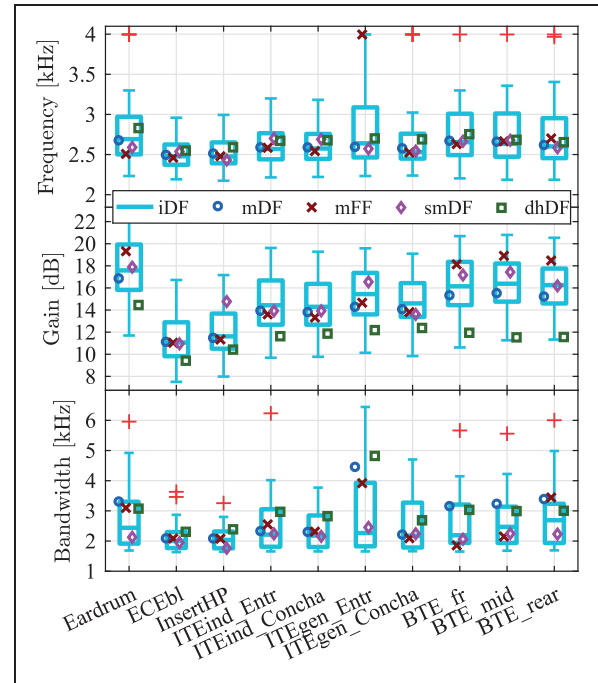
are observed in both ITEgen microphone locations, which are connected to the observed double peak in the correction functions (cf. Figure 6). Otherwise, the interindividual differences in the parameters increase with increasing distance to the eardrum.

The median resonance frequency seen in the iDF TRCF data is well conserved by the generic average functions, except for the average of dummy heads, which is systematically higher. The smallest deviation is observed in the TRCF of the ECEbl. The median gain at the peak is also generally best conserved in the mDF (iDF arithmetically averaged over subjects). With increasing distance to the eardrum, the differences between individual and average curves becomes larger, without always showing a clear tendency. At the ITEind, ITEgen, and BTE microphone locations, the gain as compared with the median of individual data is lower with the dhDF TRCF, whereas the structurally averaged DF-TRCF (smDF) yields a resonance gain that is systematically too high. The mFF (arithmetic average of individual free field TRCF) main resonance has a gain that is very comparable to the median of iDF data, except in the BTE microphone locations where it is too high (as discussed for the TRCF curves, Figure 6). The bandwidth of the main resonance is well conserved by the subject averages, as well as the structural average. At
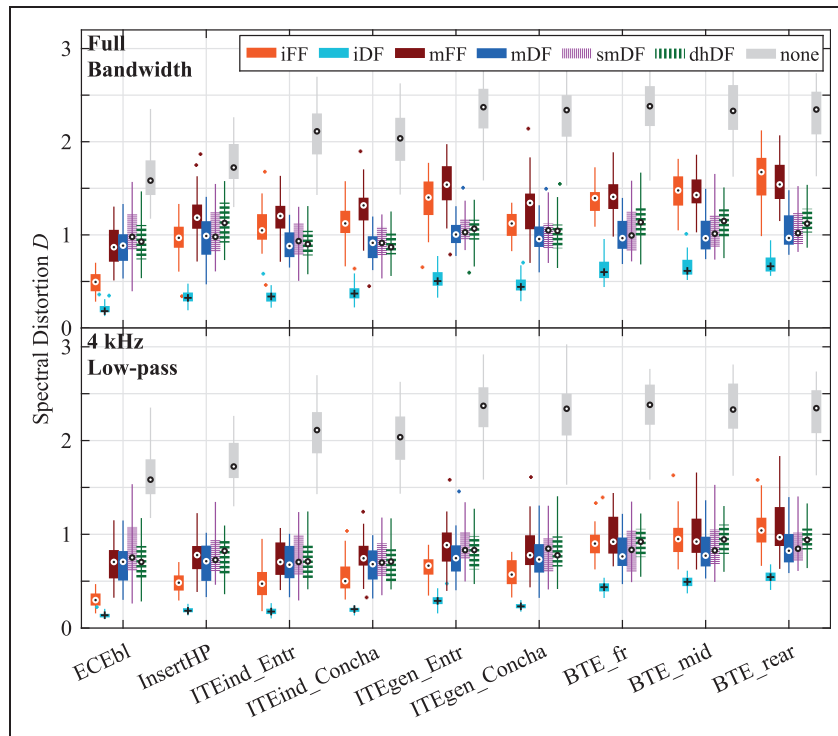
Figure 9. Spectral distortion D calculated using the model of Moore and Tan (2004), between the scene auralized with the individual HRTF to the eardrum of the open ear, and with the individual HRTFs to the hearing device microphone locations corrected by the TRCF denoted by the color. See Table 1 for an explanation of the abbreviations. The boxplots group the results in the individual subjects' ears and all seven acoustic scenes. Boxes indicate the 25%/75% quantiles, the mark in the box the median, and the vertical lines the range of results. Top panel: Full audio bandwidth. Bottom panel: Results with the stimulus low-pass filtered at 4 kHz.

the ITEgen_Entr location, the double peak in the TRCF leads to a pattern that is not reasonably usable as a correction. The bandwidth of the dhDF curve is structurally larger than the median of the iDF data.

Spectral Distortion With Corrected Hearing Device HRTFs

Figure 9 shows spectral distortion metrics D between acoustic scenes auralized with the open-eardrum HRTFs and the hearing device HRTFs corrected by the different TRCFs. The spectral distortion metric is thus an estimate of perceived distortion of natural sound transmission characteristics when a hearing device is adjusted to a target response given by the TRCFs; low values indicate a small distortion. The perceptual relevance of these distortions is further discussed in the last paragraph of this section, as well as in the Discussion section.

First, all TRCFs provide a clear benefit compared with applying no correction. For all microphone locations, the best correction (i.e., lowest residual distortion) is provided by the iDF TRCF. The performance with the iFF TRCF depends strictly on the microphone location: Whereas at the ECEbl, the residual

deviation to the optimum approaches the diffuse-field equalization performance, it rapidly increases with increasing distance from the eardrum in the other microphone locations. This increase is more pronounced with the full stimulus bandwidth, where the distortion with the iFF quickly increases above the levels of many of the generic curves. The benefit of the individual-against mean diffuse-field-TRCFs (iDF vs. mDF) decreases with increasing distance from the eardrum; the smallest difference is observed at the BTE microphone locations.

Among the generic TRCFs, the mDF (arithmetic average of iDF) correction generally yields the lowest distortion values. No substantial (and usually insignificant) difference is observed against the structural average of individual data (smDF) or the dummy head average (dhDF). The averaged free-field (mFF) correction yields very comparable values to the mDF correction when the stimuli are low-pass filtered, but in the full bandwidth condition, only at the ECEbl. At the other microphone locations, the distortion increases with the distance to the eardrum, comparable to the iFF TRCF. The modeled spectral distortion size generally decreases when applying a low-pass filter, which is more pronounced for microphone locations far away from the eardrum.
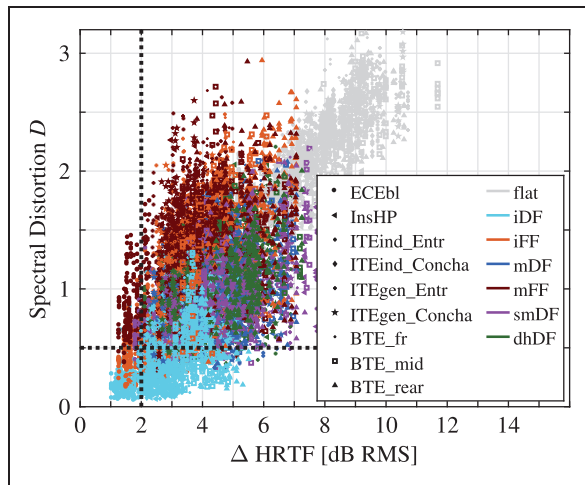
Figure 10. Spectral distortion *D* according to the model of Moore and Tan (2004) plotted against the root-mean-square (RMS) difference calculated in ERB-bins between the HRTF to the eardrum and the TRCF-corrected HRTFs to different hearing device microphones (full bandwidth only). See Table 1 for an explanation of the abbreviations. Dotted black lines illustrate the estimated thresholds of audibility: Based on a 1-dB criterion extended by a 1 dB experimental uncertainty, the spectral deviation in data points with $D \geqslant 0.5$ can be expected to be noticeable (ΔHRTF $\geqslant 2$ dB).

Figure 10 shows a comparison of the distortion values *D* from Figure 9 with the appropriate RMS spectral difference (see the Analysis Methods section for details). The comparison to a more easily interpretable physical error aims to provide a better interpretation of the dimensionless *D* value. Generally, both metrics correlate well, indicating an almost linear connection of the modeled perceptual and physical errors. In neither metric does any correction completely eliminate all errors—the deviation is always larger than about 1 dB or $D = 0.1$. As an expected noticeable difference for the spectral distortions, we thus extend the common 1 dB criterion by 1 dB of measurement uncertainty. For the vast majority of data points where $D \geqslant 0.5$, the RMS spectral difference is larger than this 2 dB boundary. Therefore, we estimate the just noticeable spectral difference to be around $D = 0.5$.

Discussion

Raw Data Compared With Previous Studies

The free- and diffuse-field-to-ear transfer functions shown in Figure 4 are generally in good agreement with previous reported studies. The most prominent differences were noted for both field conditions in the transfer functions to the ECEbl and ITE. These discrepancies are probably caused by slightly different shapes of the earplugs. The structurally lower gains in the literature

curves support the assumption that the earplugs were less deeply inserted into the ear, leading to a higher attenuation of the cavum conchae resonance (Riederer, 2004). Since the curves given by Bentler and Pavlovic are arithmetic averages of individual transfer functions, average-out effects may also have played a role. Averaging transfer functions that include peaks that are shifted in frequency leads to more gentle and lower peaks (Genuit, 1984; Mehrgardt & Mellert, 1977).

We expected the transfer function to the eardrum to be reproduced with the highest accuracy, since the reference point in the ear is very well defined. The agreement with previous measurements is excellent for the case of the diffuse sound field, whereas some differences are noted for the free field. For the free field, details of the measurement setup, such as the subjects' posture, headrests, sound source distance, and so forth, play a larger role than in the diffuse field, where these factors usually average out. One example is a ripple in the current free-field responses below 1 kHz, which is most probably caused by a comb-filter effect due to a reflection from the knees of the seated subjects, and not observed in the dummy head data. Also, average effects and average methods (see Shaw [1974] for details of the literature free-field-to-eardrum curve) could have affected the final result. The fact that the current data fits well with previous measurements that have so far been used as the standard verifies both the quality of the measurement and data-processing procedures employed. It also validates the averaging procedure across directions to approximate diffuse-field responses. The data therefore enables us to utilize the current data, for example, to estimate the expected individual deviations from the average and the expected hearing device style-specific deviations. For consistency, only correction curves from data observed in this study are further evaluated, although similar results would be expected when using the curves given by Bentler and Pavlovic (1989, 1992).

The fact that the differences between subjects are larger in the free-field transfer functions can also be explained by the finer details in the free-field transfer functions, which are averaged out in the diffuse field. In both cases, individual differences increased with decreasing distance from the eardrum. This results from the simple fact that an increasing number of acoustic transmission elements that are individual to each person are captured, thus yielding more individual features in the transfer function.

Corrections Functions Related to External Ear Acoustics

The TRCF should artificially restore features that would occur during sound transmission from the respective microphone location to the eardrum of an open ear.

This comprises directionally independent resonances, such as cavity resonances of the ear canal and the cavum conchae, as well as directionally dependent spatial features (Shaw & Teranishi, 1968). These cues can best be inspected in the directionally resolved RTFs shown in Figure 5, which also contain the individual TRCFs for the free field (frontal incidence, red curves) and diffuse field (blue curves).

The full HRTF is shown for the eardrum, where all acoustic cues generated by the external ear are included in the transfer functions. Starting at frequencies of several hundred Hz and more prominent for frequencies > 1 kHz, there is a considerable spread over incidence directions, prominently between ipsi- and contralateral incidence. The physical correlate is the head-shadowing effect. This variation is, to a large extent, cancelled out in the RTFs in most of the hearing device microphone locations, with the exception of the BTE loctions. We conclude that in all hearing devices where the microphone is located in the concha, head-shadowing effects, and therefore interaural level cues, are conserved to a very high extent (consistent with Kayser et al., 2009).

The RTF from the ECEbl to the eardrum shows no considerable dependence on the incidence direction, and structurally consists of several peaks and notches that can be recognized in both subjects. The variance across incidence directions increases gradually with frequency and appears to be random, probably originating from experimental uncertainty. We conclude that sound transmission through the ear canal is not directionally dependent, verifying previous studies (Algazi, Avendano, & Thompson, 1999; Hammershøi & Møller, 1996; Mehrgardt & Mellert, 1977). The lowest resonance in the ECEbl to eardrum RTF corresponds to the $\lambda/4$ resonance of the ear canal. Features at higher frequencies do not fit with a simple transmission-tube model and probably depend on the eardrum impedance and the shape of the ear canal (c.f. discussion of Hammershøi & Møller, 1996).

Whereas at the ECEbl, the RTF decreases after the first peak, the peak extends further to higher frequencies in most other microphone locations. This can be understood as a superposition of the $\lambda/4$ resonance of the ear canal and broader peaks at higher frequencies. The physical correlate of this directionally independent feature is the cavity resonance of the cavum conchae (Shaw & Teranishi, 1968), which is increasingly attenuated through filling by the hearing device. These resonances are best assessed for the diffuse-field RTFs. The InsertHP only partly fills the concha, and only higher frequencies of the resonance are affected. At low frequencies, a shape that is much comparable to the ECEbl is observed. By contrast, towards higher frequencies (> 4 kHz), a larger correction is necessary to restore the transfer function to the eardrum. The more

the device fills up the cavum conchae or the further the microphone is positioned off the ear canal entrance, the less of the concha resonance is captured by the microphone. Consequently, more gain in the corresponding frequencies has to be applied in the TRCF. The ITEind fills the concha very uniformly, consequently the resonance is attenuated in all of its frequencies, leading to larger and smoother amplitude curves of the RTF in the diffuse field. For the ITEgen device, the concha is filled less uniformly and the microphone protrudes further, leading to more peaky RTFs. The reasons are probably cavity resonances in the residual cavum conchae segments.

In the InsertHP and ITE microphone locations, the variance across directions increases rapidly above approximately 4 kHz. At this frequency, the wavelength becomes comparable to the size of the pinna and concha. Therefore, spectral characteristics created by these structures are biased by the microphone placement and obstruction of the pinna in these conditions. The result excellently reproduces previous data on comparable hearing device microphones (Durin et al., 2014; Hoffmann et al., 2013b). In the band between 4 and 8 kHz in the RTFs for free-field (i.e., frontal) incidence, a distinct notch is observed. Referring to Figure 4, this notch is present in the free-field-to-eardrum and ear canal entrance transfer function, but not in the ITE condition. This behavior can be related to the restoration of a destructive interference of a wave that directly enters the ear canal, and a component that is reflected by the concha back wall, often referred to as *Concha Notch* (Butler & Belendiuk, 1977). This notch occurs in the median plane in frontal directions, but not for other incidence directions. The same observation was made in the RTFs of the BTE locations, but there, the concha notch is overlaid by other structures and is less prominent than in the ITE and InsertHP locations. Given the average over the free-field correction functions (TRCF-mFF, see Figure 6), no large difference for this spectral region is observed between the InsertHP and the ITE locations, particularly the ITEind. Therefore, the spectral notch is not better included in the InsertHP, although it obstructs the cavum conchae less than the earmold of the ITEind. However, the conservation of this feature in the different microphone locations is subject to large differences between the individual ears, as can be observed in the individual RTFs given in Figure 5.

Free-Field Versus Diffuse-Field Correction

The findings with the spectral distortion model shown in Figure 9 indicate a preference for correcting the target of a hearing device to the diffuse field rather than to the free field (i.e., frontal incidence). The results were obtained with a set of realistic acoustic scenes with several desired and distracting sound sources and reverberation

in a complex spatial setup, and are therefore expected to hold in real-life acoustic environments.

The analysis in the previous section revealed that several directional cues for the prominent frontal direction (like the Concha Notch) can only be included in the sound field if it is included in the TRCF. Such a procedure might, on the one hand, lead to a better spatial reproduction of the environment due to the inclusion of such features, but on the other hand, this will lead to larger errors for other directions than when applying a response correction that is correct for the directional average. In daily life, neither of the extreme cases, free field and diffuse field, occurs exactly. If the correction aims at equalizing for one prominent direction, the frontal direction is surely a well-grounded choice. However, due to reverberation, the pressure created by a sound source in the viewing direction is normally very different from what would be created in the free field. In rooms, fine structures, particularly notches, are flattened by reflections from objects and by reverberation (Shinn-Cunningham, Kopco, & Martin, 2005). This makes the transfer function to the eardrum generally more alike to that observed in an ideal diffuse field. Similar considerations and conclusions were drawn by Killion and Monsor (1980), but without a detailed assessment. The diffuse field may, however, not be the optimal reference for defining the TRCF. Whether a (weighted) average of the RTF over a constrained set of directions is a better option would be an interesting question for future psychoacoustic experiments.

Besides the directional effects originating from the pinna, a distinct ripple was seen in the free-field TRCF below 2 kHz, but not in the diffuse-field TRCF. The ripple most probably originates from a reflection from the knees of the subjects that only occurs for incidence directions in the frontal median plane. Given the size of the effect, large influences on the results regarding spectral distortions are not expected, besides the fact that this feature could also be regarded as a relevant directional cue.

Altogether, we recommend that the hearing device frequency response should be equalized by a correction that is correct to the diffuse field rather than to the free field. This holds for both when individual measurements are available, and for when generic correction functions are used. Future research should examine whether even better results can be achieved by correcting the response for the (weighted) average of different incidence directions, instead of the extreme cases of the free or diffuse field.

Individual versus Generic Correction and Averaging Methods

Given the notable variations of the TRCFs between individual ears shown in Figures 6 and 7, it seems obvious

to assume that using individual TRCFs is preferable against a generic correction derived from averaging subject data or dummy head measurements. This benefit was verified by the spectral distortion model results from Figure 9 for the diffuse-field correction curves (iDF vs. mDF) in general, whereas for free-field corrections, the benefit depends on the microphone location in the ear. For the ECEbl, the directional information is valid over a wide frequency range, and the iFF TRCF provides a benefit against generic TRCFs. For the other microphone locations, the spectral distortion with the iFF TRCF was larger than with generic diffuse field (mDF, smDF, and dhDF) TRCFs. When the stimuli are low-pass filtered, the spectral distortion with the iFF was, for all microphone locations, at least as small as with the generic diffuse-field TRCF. Apparently, there is a trade-off between conserving the main resonance as the most characteristic individual feature and other errors that occur in different directions, mostly in the higher frequency regime above 4 kHz. Thus, whereas both the main resonance and the average high-frequency behavior of the iDF-TRCF fit all incidence directions on average, the notch in the higher frequencies of the iFF curve leads to higher errors for incidence directions other than from the front. This explanation fits well with the observed dependence on the stimulus bandwidth: When a 4-kHz low pass is applied, the correct spatial information is captured by all microphone locations except the BTE device. Consequently, an iFF correction is only beneficial with respect to the generic TRCFs with the low-pass stimulus.

As a general statement, the benefit of the individual TRCF compared with average values decreased with increasing distance to the eardrum, and almost vanishes at the BTE locations. This is rather surprising, because the further away from the eardrum the hearing device microphone is located, the more acoustic features have to be included in the TRCF—features that are in principle individual to each ear. Apparently, the superposition of many of such features leads to a decrease of individuality, particularly in the diffuse-field correction curves. This conclusion is even stronger when the bandwidth is restricted.

Among the generic DF correction functions (mDF, smDF, and dhDF), no large differences in spectral distortion were observed. Generally, the arithmetic average over individual TRCFs (mDF) produces the least spectral distortions, and the ranking of structural average and dummy-head average depends on the microphone location. The difference between the mDF and dhDF-TRCFs is caused by the differences in geometry and eardrum impedance between the artificial and real ears. Both contain a comparable degree of spectral detail, but the dhDF lies 2 to 5 dB lower in level around the main resonance. The level offset might be

explained by a mismatch between the impedances of the ear simulators (IEC711) and the real eardrum impedances in the subjects. Since the level gap increases with the distance from the eardrum, a slight misalignment in the dimensions of the pinna or the acoustic impedance of its material might also explain the mismatch. However, the spectral distortion observed with the correction function obtained from the dummy heads is only slightly higher than with the subject average (Figure 9). We conclude that when no individual data are available, using the TRCF obtained from dummy head measurements is justified by the current results. This holds especially when considering the effort that is connected to measurements like the ones presented here, compared with the same experiments with dummy heads. Based on the current results, it might be beneficial to increase the gain of the main resonance of the dummy head TRCF as a heuristic correction.

The structural mean of the subjects' diffuse field TRCFs (smDF) avoids smearing effects due to arithmetic averaging over shifted peaks. The result should be a "typical" curve, which would be represented in a better conservation of structural parameters of the TRCFs than with a standard averaging procedure. However, the structural parameters of the main resonance are not better represented by this structural average when compared with the arithmetic mean. Whereas the resonance frequency conserves well the subject median in all microphone locations, the gain is overestimated in the structural mean (see Figures 6 and 8). Only the bandwidth of the main resonance is better conserved with the structural average. Given these results and the fact that the TRCF-smDF does not provide a benefit compared with the standard average in HRTF correction in terms of spectral distortion, we conclude that this simple structural model average should not be used. More sophisticated models may provide a better fit. This, however, makes averaging across the model parameters more challenging, and the model requires additional knowledge such as the geometric ear dimensions (Genuit, 1984).

We conclude that an arithmetic average of individual human subjects' diffuse-field TRCF is the best option to obtain generic TRCFs. However, a correction curve based on dummy head data can be used with almost the same performance and could be improved by a heuristic correction.

Perceptual Relevance of Spectral Distortions

The spectral distortion model used was based on physiologically motivated excitation patterns on the basilar membrane and yields a good prediction when a spectral difference is perceivable between two arbitrary stimuli (Moore & Tan, 2004). This prediction worked very well for spectral naturalness judgments of electroacoustic transducers. However, in this context, it might be hypothesized that a spectral deviation from the individually correct external ear transfer function through which the subjects hear the world every day may be perceived as more disturbing than the same spectral distortion to the source signal. Final judgments about the benefit of individual response correction functions can thus only be made based on subjective listening experiments. Based on a common 1-dB criterion extended by the experimental uncertainty, the just noticeable spectral distortion was estimated to be around $D = 0.5$ (see Figure 10 and Results section).

This boundary is exceeded in the majority of conditions, and therefore the spectral distortions can, in general, be considered as relevant. Only when the iDF TRCF is applied did a considerable share of data points fall below this boundary. All these conditions belong to microphone locations inside the pinna, and include all devices except the BTE. The defined threshold was derived from the appropriate just-noticeable difference in psychoacoustic A-B comparisons, and therefore the lower boundary of detection sensitivity. In real-world environments with dynamically changing sound sources, the threshold for noticeable differences is likely to be much higher.

Altogether, we conclude that the construction of hearing devices that are acoustically transparent is possible in the perceptual sense with all device styles except the BTE. On the other hand, the result shows that acoustic transparency can only be achieved using individualized target-correction functions.

Conclusions

We studied external ear acoustics related to the task of equalizing hearing devices to the acoustics of the ear of the individual subject. In particular, correction functions that transform the pressure at the device's microphone to the reference observed at the open eardrum (here referred to as TRCF) were derived using different approaches and for various hearing device styles. For that purpose, we measured HRTFs to a comprehensive set of hearing device microphone locations and the eardrum in 16 human subjects and 2 dummy heads from 91 incidence directions. The HRTF dataset and the derived TRCFs are publicly available.[1]

TRCFs were calculated for each individual ear based on the relative transfer function between the microphone location and the eardrum of the open ear, for both a free- and diffuse-sound field. Generic correction functions that do not include measurements on the specific subject were derived through arithmetic and structural averaging of the individual data, as well as using data from commercial dummy heads with standardized IEC711 ear

simulators. Our main conclusions regarding the correction functions are:

- The TRCF depends greatly on the exact location of the microphone in the ear, as well as the degree to which the ear is filled by the device. It restores acoustic features that would normally occur during sound transmission between the microphone location and the eardrum of the open ear, most importantly a directionally independent resonance of the ear canal and cavum conchae between approximately 1 and 7 kHz.
- Diffuse-field equalization of the microphone signal to the open-ear reference should be preferred over free-field equalization.
- The TRCF differs between individuals, like each ear is different. Structural similarities between individuals in each microphone location exist, but these features are shifted in frequency and level.
- Individualized TRCFs provide a significantly better adaptation to the individual ear acoustics than generic correction functions. Using individual diffuse-field correction, it is possible to equalize the target of most hearing device styles down to an error compared with open-ear listening that is probably not noticeable in common acoustic scenes. Generic correction functions lead to spectral distortions that are probably noticeable.
- Among generic equalization functions, the best result is achieved when using an arithmetic average of the individual diffuse-field TRCFs. Using data obtained from dummy head measurements is a valid alternative that creates only marginally larger errors. Only minor differences to averaging over subjects were observed in the present data, most prominently a systematically lowered TRCF gain. A structural mean calculated using a rudimentary two-resonator model of the external ear resulted in a slightly poorer result than the arithmetic average.
- The benefit of diffuse-field correction and individual transfer functions interact with the device style: The less the hearing device microphone captures spatially dependent cues correctly, the more a diffuse-field TRCF should be preferred over a free-field TRCF. The benefit of individual TRCFs decreases with increasing distance of the hearing device microphone from the eardrum and increasing filling of the cavum conchae, particularly if it is filled uniformly. This means that individual correction functions are most beneficial in devices in the ear canal (such as CIC hearing aids) and least beneficial in BTE devices.

The results demonstrate the benefit of individualized hearing device fitting using probe tube measurements.

According to our data, probe tube measurements should be performed in (approximated) diffuse-field conditions, for example, using several loudspeakers in a reverberant room. On the other hand, the generic TRCFs can be directly applied in devices that utilize a nonindividualized response target, such as consumer hear-through headsets (Hoffmann et al., 2013a).

For constructing acoustically transparent hearing devices, the optimum style would allow microphone positioning at the earcanal entrance capturing the full spatial HRTF information. Our data once more confirms that full spatial acoustic cues can only be observed with a microphone in the ear canal (Algazi et al., 1999; Hammershøi & Møller, 1996), and thus acoustic transparency in the strict physical sense is only achievable with devices in the ear canal and individualized target responses. However, space constraints and nonindividualized shells require filling at least a part of the cavum conchae in many applications—in this case, the results indicate that it is beneficial to occlude the concha as uniformly as possible, even if it means that the device becomes larger. This holds especially if generic correction functions are utilized. With devices that have the microphone located inside the cavum conchae and using individualized TRCF, we conclude that it is possible to achieve acoustic transparency in a perceptual sense. With BTE microphones, it does not seem possible to construct acoustically transparent devices.

Future work should include a psychoacoustic validation of the findings. Open questions are to what extent noticeable differences to the usual transfer function to the eardrum translate to an impaired sound quality, and, specifically, what the subjective benefit of individual correction functions really is.

Note

1. The HRTF database and the TRCFs derived are available for download at http://medi.uni-oldenburg.de/hearing devicehrtfs/.

ORCID iD

Florian Denk http://orcid.org/0000-0003-3490-123X

References

Algazi, V. R., Avendano, C., & Thompson, D. (1999). Dependence of subject and measurement position in binaural signal acquisition. *Journal of the Audio Engineering Society*, *47*(11), 937–947.

Bentler, R. A., & Pavlovic, C. V. (1989). Transfer functions and correction factors used in hearing aid evaluation and research. *Ear and Hearing*, *10*(1), 58–63. doi:10.1097/00003446-198902000-00010

Bentler, R. A., & Pavlovic, C. V. (1992). Addendum to "Transfer Functions and Correction Factors Used in Hearing Aid Evaluation and Research". *Ear and Hearing*, *13*(4), 284–286. doi:10.1097/00003446-199208000-00012

Breebart, J., & Kohlrausch, A. (2001). *Perceptual (ir)relevance of HRTF magnitude and phase spectra*. Paper presented at the Audio Engineering Society Convention, Amsterdam, Netherlands, 110, 1–9. New York, NY: Audio Engineering Society.

Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *The Journal of the Acoustical Society of America*, *106*(3), 1465–1479. doi:10.1121/1.427180

Butler, R. A., & Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *The Journal of the Acoustical Society of America*, *61*(5), 1264–1269. doi:10.1121/1.381427

Denk, F., Heeren, J., Ewert, S. D., Kollmeier, B., & Ernst, S. M. A. (2017). *Controlling the head position during individual HRTF measurements and its effect on accuracy*. In proceedings of Fortschritte der Akustik—DAGA, Kiel, Germany.

Denk, F., Hiipakka, M., Kollmeier, B., & Ernst, S. M. A. (2017). An individualised acoustically transparent earpiece for hearing devices. *International Journal of Audiology*. Advance online publication. doi:10.1080/14992027.2017.1294768

Denk, F., Kollmeier, B., & Ewert, S. (2018). Removing reflections in semianechoic impulse responses by frequency-dependent truncation. *Journal of the Audio Engineering Society*, *66*(3), 146–153. doi:10.17743/jaes.2018.0002

Dillon, H. (2012). *Hearing Aids* (2nd ed.). Turramurra, Australia: Boomerang Press.

Durin, V., Carlile, S., Guillon, P., Best, V., & Kalluri, S. (2014). Acoustic analysis of the directional information captured by five different hearing aid styles. *The Journal of the Acoustical Society of America*, *136*(2), 818–828. doi:10.1121/1.4883372

Genuit, K. (1984). *Ein Modell zur Beschreibung von Aussenohrübertragungseigenschaften* [A model for describing the external ear transmission properties] (PhD Thesis). Aachen, Germany: RWTH Aachen.

Grimm, G., Kollmeier, B., & Hohmann, V. (2016). Spatial acoustic scenarios in multichannel loudspeaker systems for hearing aid evaluation. *Journal of the American Academy of Audiology*, *27*(7), 557–566. doi:10.3766/jaaa.15095

Hammershøi, D., & Møller, H. (1996). Sound transmission to and within the human ear canal. *The Journal of the Acoustical Society of America*, *100*(1), 408–427. doi:10.1121/1.415856

Härmä, A., Jakka, J., Tikander, M., Karjalainen, M., Lokki, T., Hiipakka, J., & Lorho, G. (2004). Augmented reality audio for mobile and wearable appliances. *Journal of the Audio Engineering Society*, *52*(6), 618–639.

Hellstrom, P.-A., & Axelsson, A. (1993). Miniature microphone probe tube measurements in the external auditory canal. *The Journal of the Acoustical Society of America*, *93*(2), 907–919. doi:10.1121/1.405452

Hirahara, T., Sagara, H., Toshima, I., & Otani, M. (2010). Head movement during head-related transfer function measurements. *Acoustical Science and Technology*, *31*(2), 165–171. doi: 10.1250/ast.31.165.

Hoffmann, P., Christensen, F., & Hammershøi, D. (2013a). *Insert earphone calibration for hear-through options*. Paper presented at the audio engineering society conference: 51st international conference: Loudspeakers and headphones. Helsinki, Finland: Audio Engineering Society.

Hoffmann, P., Christensen, F., & Hammershøi, D. (2013b). Quantitative assessment of spatial sound distortion by the semi-ideal recording point of a hear-through device. *Proceedings of Meetings on Acoustics, Acoustical Society of America*, *19*, 050018.

Kayser, H., Ewert, S., Anemüller, J., Rohdenburg, T., Hohmann, V., & Kollmeier, B. (2009). Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing*, *2009*(1), 298605. doi:10.1155/2009/298605

Killion, M. C. (1979). *Design and evaluation of high-fidelity hearing aids* (PhD Thesis). Northwestern University, Evanston, IL.

Killion, M. C., Berger, E. H., & Nuss, R. A. (1987). Diffuse field response of the ear. *The Journal of the Acoustical Society of America*, *81*(S1), S75–S75. doi:10.1121/1.2024388

Killion, M. C., & Monsor, E. L. (1980). CORFIG: Coupler response for flat insertion gain. In G. A. Studebaker, & I. Hochberg (Eds.), *Acoustical factors affecting hearing aid performance* (1st ed., pp. 149–168). Baltimore, MD: University Park Press.

Kuhn, G. F. (1979). The pressure transformation from a diffuse sound field to the external ear and to the body and head surface. *The Journal of the Acoustical Society of America*, *65*(4), 991–1000. doi:10.1121/1.382606

Kuhn, G. F., & Burnett, E. D. (1977). Acoustic pressure field alongside a manikin's head with a view towards in situ hearing-aid tests. *The Journal of the Acoustical Society of America*, *62*(2), 416–423. doi:10.1121/1.381516

Lindau, A., & Brinkmann, F. (2012). Perceptual evaluation of headphone compensation in binaural synthesis based on non-individual recordings. *Journal of the Audio Engineering Society*, 60(1/2), 54–62.

Madaffari, P. L. (1974). Pressure variation about the ear. *The Journal of the Acoustical Society of America*, 56(S1), S3–S3. doi:10.1121/1.1914163

Majdak, P., Balazs, P., & Laback, B. (2007). Multiple exponential sweep method for fast measurement of head-related transfer functions. *Journal of the Audio Engineering Society*, 55(7/8), 623–637.

Mehrgardt, S., & Mellert, V. (1977). Transformation characteristics of the external human ear. *The Journal of the Acoustical Society of America*, 61(6), 1567–1576. doi:10.1121/1.381470

Moore, B. C., & Tan, C.-T. (2004). Development and validation of a method for predicting the perceived naturalness of sounds subjected to spectral distortion. *Journal of the Audio Engineering Society*, 52(9), 900–914.

Mueller, H. G. (2001). Probe microphone measurements: 20 years of progress. *Trends in Amplification*, 5(2), 35–68.

Orfanidis, S. J. (1997). Digital parametric equalizer design with prescribed Nyquist-frequency gain. *Journal of the Audio Engineering Society*, 45(6), 444–455.

Rämö, J., & Välimäki, V. (2012). Digital augmented reality audio headset. *Journal of Electrical and Computer Engineering*, 2012, 1–13. doi:10.1155/2012/457374

Riederer, K. A. (2004). Part IVa: Effect of cavum conchae blockage on human head-related transfer functions. *Proceedings of 18th International Congress on Acoustics ICA*, Kyoto, Japan.

Shaw, E. A. G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *The Journal of the Acoustical Society of America*, 56(6), 1848–1861. doi:10.1121/1.1903522

Shaw, E. A. G. (1980). The acoustics of the external ear. In G. A. Studebaker, & I. Hochberg (Eds.), *Acoustical factors affecting hearing aid performance*. Baltimore, MD: University Park Press.

Shaw, E. A. G., & Teranishi, R. (1968). Sound pressure generated in an external ear replica and real human ears by a nearby point source. *The Journal of the Acoustical Society of America*, 44(1), 240–249. doi:10.1121/1.1911059

Shaw, E. A. G., & Vaillancourt, M. M. (1985). Transformation of sound pressure level from the free field to the eardrum presented in numerical form. *The Journal of the Acoustical Society of America*, 78(3), 1120–1123. doi:10.1121/1.393035

Shinn-Cunningham, B. G., Kopco, N., & Martin, T. J. (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses. *The Journal of the Acoustical Society of America*, 117(5), 3100–3115. doi:10.1121/1.1872572

Wiener, F. M., & Ross, D. A. (1946). The pressure distribution in the auditory canal in a progressive sound field. *The Journal of the Acoustical Society of America*, 18(2), 401–408. doi:10.1121/1.1916378