DOI: 10.61782/fa.2023.1104

forum acusticum 2023

SPECTRAL AND PSYCHOACOUSTIC EVALUATION OF HEAD-RELATED TRANSFER FUNCTIONS CALCULATED AT THE BLOCKED EAR CANAL AND THE EARDRUM

Katharina Pollack[1*] **Fabio di Giusto**[2,3] **Daniel Sinev**[4,5]
Piotr Majdak[1]

[1] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
[2] Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300 B, B-3001, Heverlee, Belgium
[3] Flanders Make@KU Leuven
[4] Sonova Consumer Hearing GmbH, Wedemark, Germany
[5] Institut für Kommunikationstechnik, Leibniz Universität Hannover, Germany

ABSTRACT

Head-related transfer functions (HRTFs) are essential for personalised binaural audio reproduction and have been defined at the blocked ear canal for reasons originating in acoustic measurements. With the possibility to numerically calculate HRTFs, they can be investigated at the eardrum, given an accurate representation of the geometry of the ear canal. In our study, we evaluated the contribution of the ear canal to the sound-localisation performance in the median plane. We used meshes of head and pinnae including ear canals to calculate HRTFs at both the blocked ear canal and the eardrum. The resulting HRTFs showed spectral differences of up to 25 dB in frequency regions above 3 kHz. Simulations based on directional transfer functions (DTFs) obtained with an auditory model for the sagittal-plane sound localisation showed negligible differences in the predicted localisation performance. Based on these results, the applicability of numerically calculated HRTFs for binaural audio reproduction and sound localisation experiments is discussed.

Keywords: *head-related transfer functions, blocked ear canal, eardrum, numerical calculation, auditory modeling*

**Corresponding author*: *katharina.pollack@oeaw.ac.at.*

1. INTRODUCTION

Head-related transfer functions (HRTFs) describe the acoustic filtering due to a listener's individual anatomy and are essential for plausible binaural audio playback [1]. Previously, the microphone placement in an acoustic measurement of HRTFs has been investigated and discussed [2, 3], and negligible sound-pressure differences between placing the microphone at the entrance of the ear canal and 1 mm inserted into the ear canal were found for frequencies between 3 kHz and 16 kHz. Recently, a database of ten listener-specific meshes including the ear canal up to the eardrum has been published [4], which served as motivation for this study. Since HRTFs can be calculated quickly using simulation tools, we aim at reproducing the findings of [2] and compare the HRTFs calculated at the blocked ear canal and at the eardrum.

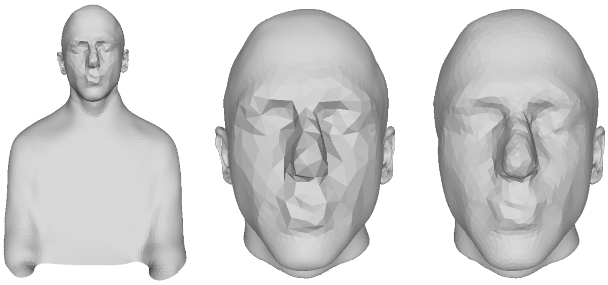
This paper is part of a three-paper study including measurements and simulations of a pinna replicas of a subject from [4], which are presented in [5, 6].

2. METHODS

2.1 IHA Database

The recently published IHA database contains human geometry scans including torso, head and pinnae up to the eardrum [4]. The geometries from this database can be used to calculate HRTFs at the blocked ear canal and the eardrum. The torsi of the ten subjects were removed, because only a frequency range between 3 and

20 kHz was considered, and because HRTFs for a mesh including shoulders and torso are costly to calculate albeit their influence is negligible in this frequency range. In the case of calculations based on a blocked ear canal, the ear canal was removed in the animation software Blender v3.1.4 [7].Then, the meshes were graded using a curvature-adaptive approach such that the highest resolution is provided at the ipsilateral pinna, and the resolution decreases towards the contralateral pinna [8] [1, 2]. Figure 1c shows the geometry modification that was applied to the ten available meshes. Figure 2 shows the graded meshes and virtual microphone position of a subject-specific geometry including the ear canal up to the eardrum, and with a blocked ear canal.



(a) (b) (c)

Figure 1: IHA09 geometry: **(a)** raw (incl. torso), **(b)** without head, and **(c)** remeshed for the calculation of the left-ear HRTF.

2.2 HRTF calculation

HRTFs were calculated on the graded meshes using Mesh2HRTF v1.0 [9, 10] [3] and stored as SOFA files using the SOFA Toolbox v2.1 [11] [4]. We used Mesh2HRTF because it yields a maximum error of 0.5 dB for 1 kHz with respect to analytical solutions when comparing the pressure distribution on a sphere [12], and because it also yields similar performance in simulated localisation experiments by means of sound-localisation errors [13]. We additionally compared Mesh2HRTF's underlying boundary element method (BEM) computation to an analytical solution because the frequency range of interest in this investigation was between 3 and 20 kHz. A head in the centre of the coordinate system was approximated by a graded sphere with 0.1 m radius using the previously mentioned mesh grading tool and a velocity source at the position of the left ear, i.e., an element closest to the positive interaural axis. The evaluation grid was a uniformly sampled sphere with 1.2 m radius and 2,556 positions around the head for which the complex pressure was calculated.

(a) (b)

Figure 2: Geometry and the microphone placement (dot) **(a)** with the ear canal (up to the eardrum) and **(b)** with the blocked ear canal.

2.3 Evaluation

On the acoustic level, HRTFs were compared using the spectral difference

$$SD_{ij} = 20 \cdot log_{10}\left(\frac{|(H_{ij}|}{|H_{ij,ref}|}\right), \qquad (1)$$

where $|H|$ denotes the absolute value of the complex pressure of a calculated HRTF for 1,550 positions on a sphere with 1.2m radius, $i = 1, \ldots, 128$ denotes the ith frequency bin (linearly spaced), $j = 1, \cdots, 10$ denotes the jth subject ID from the database. H_{ij} denotes the HRTF calculated at the blocked ear canal, and ref denotes the reference HRTF, i.e., calculated at the eardrum.

On the psychoacoustic level, we used the model of the sagittal-plane sound localisation implemented as `baumgartner2014` [14] in the auditory modeling toolbox (AMT, [15]) [5] to predict the quadrant error rate and the polar error in the median plane. Before being put in the auditory model, all HRTFs were diffuse-field compensated, i.e., converted to directional transfer functions

[1] https://github.com/cg-tub/hrtf_mesh_grading/
[2] https://mesh2hrtf-tools.sourceforge.io/
[3] https://mesh2hrtf.org/
[4] https://github.com/sofacoustics/SOFAtoolbox
[5] http://amtoolbox.org/

(DTFs) by `SOFAhrtf2dtf.m`, and transformed into the time domain using `SOFAconvertConventions.m` from the SOFA Toolbox. We applied the model using the DTFs at the eardrum as reference, which we interpreted as "natural" listening setting, i.e., a person used to DTFs at their eardrum is presented with DTFs at their blocked ear canal.

3. RESULTS AND DISCUSSION

Figure 3 shows the outcome of a BEM computation on a sphere with 1.2 m radius and its difference to the analytical solution. Only a quarter of the horizontal plane is plotted, because the geometry was symmetric by the median and frontal plane. The magnitude error was less than 1 dB for all frequencies on the ipsilateral side, indeed. On the contralateral side, the magnitude error was up to 6 dB due to the coarser grading, which is not relevant for this study, because the contralateral ear has negligible contribution to sound-localisation performance [16].

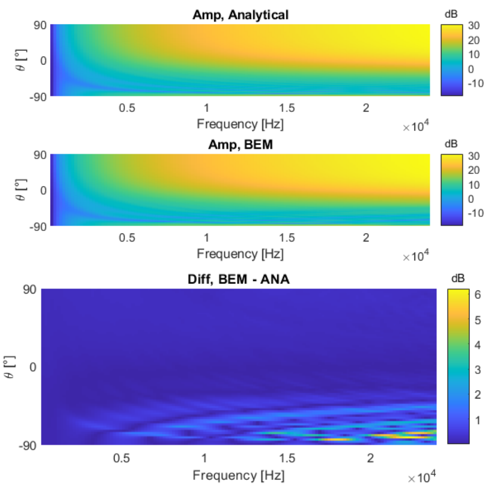
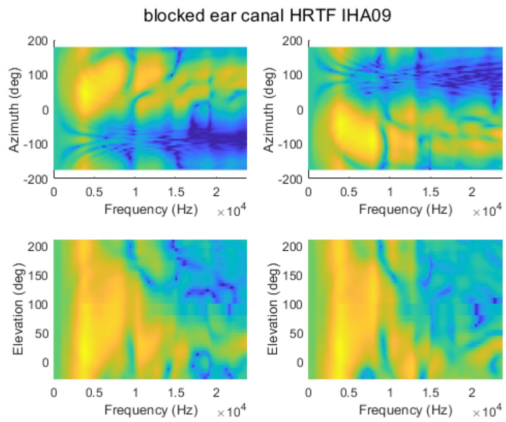


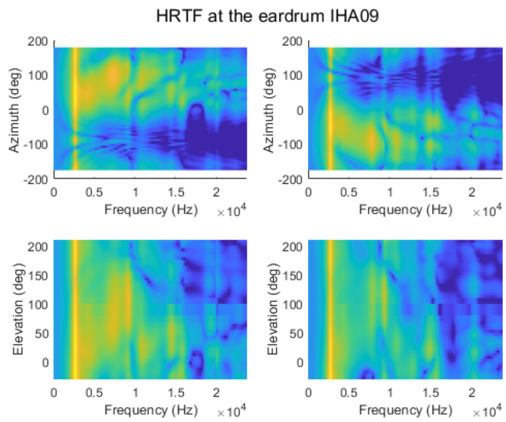
Figure 3: Sound pressure (in dB) resulting from the analytical solution (top), from the BEM calculation (center), and the difference in their magnitude (bottom).

Figures 4 and 5 show the calculated HRTFs and DTFs in the horizontal and median plane. In the HRTFs, the effect of the ear canal is apparent for all sound directions, supported by a prominent, almost direction-independent resonance at about 2.8 kHz, and a high attenuation of higher frequencies, especially at the contralateral side. In the DTFs, there is little spectral difference between the microphone placement at the blocked ear canal and at the eardrum.

Figure 6 shows the spectral differences calculated between the HRTFs at the blocked ear canal and at the eardrum as in Eq. 1. The biggest difference was the peak resonance of the ear canal at around 2.8 kHz and its third harmonic at around 8.4 kHz, which seem to be direction independent. There seems to be a more complex error pattern for frequencies higher than 13 kHz. Within the frequency range of 15 to 20 kHz, the largest magnitude errors for both ears were approximately 25 dB.

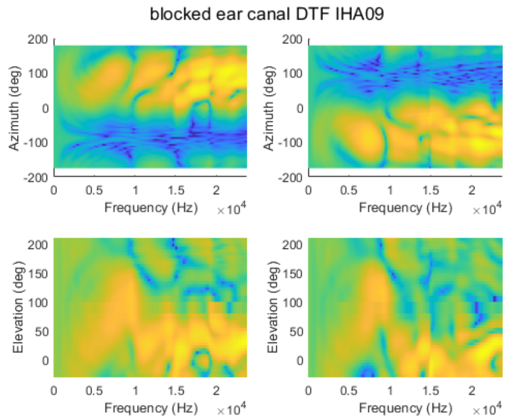


(a)

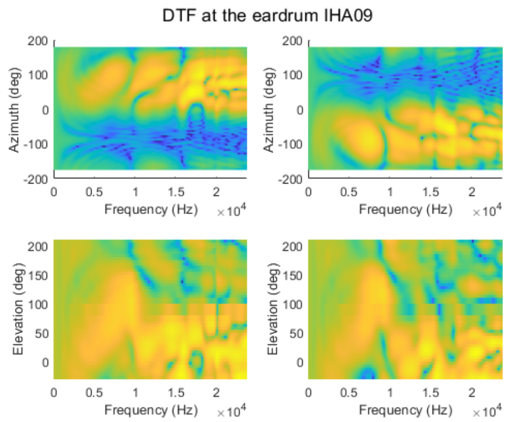


(b)

Figure 4: Spectral magnitude of the HRTFs (in dB) calculated **(a)** at the blocked ear canal and **(b)** at the eardrum.



(a)



(b)

Figure 5: Spectral magnitude of the DTFs (in dB) calculated **(a)** at the blocked ear canal and **(b)** at the eardrum.

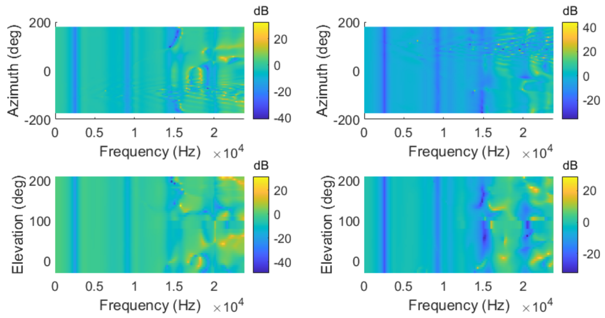


Figure 6: Spectral differences (in dB) between the HRTFs calculated at the eardrum and the blocked ear canal. The left ear (left column) and the right ear (right column) for the horizontal plane (top row) and the median plane (bottom row).

Figure 7 shows the reproduced Fig. 6a from [2], the standard deviations of calculated pressure magnitude at the blocked ear canal and the eardrum for all ten subjects in the database and the average of the ten subjects. One difference between these two datasets is that the initial one from 1989 is based on measured data of 6 subjects at 356 source positions, and in this investigation, we included calculations of ten subjects at 1,550 source positions. Another difference is that due to the nature of the measurement procedure, the microphone was inserted only about 10 mm into the ear canal, and the calculation was evaluated at the position of the eardrum. In this investigation, the standard deviations for frequencies higher than 11 kHz were between -25 and 25 dB. These results confirm a high direction-dependent directional spread for frequencies above 10 kHz consistent with [3].

Figure 8 shows the change in predicted sound-localisation errors for calculated DTFs, i.e., quadrant error (QE) rate and polar error (PE) for all ten subjects. IHA05 shows the largest deviation, their QE rate changed by 3%, and their PE changed by 1.5°. Literature suggests that errors of this size may be perceivable [17], however, the error results for other listeners mostly do not change more than 1% and 1°. To this end, the model input is limited by the utilisation of DTFs, and thus, it is difficult to estimate a sound-localisation performance using HRTFs. This suggests that acoustic experiments with human subjects might be necessary to investigate the influence of the ear canal further.

4. CONCLUSIONS

In this investigation, we exploited a new database of geometries of ten subjects including their ear canal up to the eardrum. We calculated HRTFs and DTFs from these geometries considering two conditions: at the eardrum and at the blocked ear canal. The latter condition reflects a typical state-of-the-art acoustical HRTF measurement condition. Our results show spectral differences of up to ±25 dB for frequencies higher than 10 kHz in both ears. These differences were drastically reduced when the direction-independent components of the HRTFs were removed, which was done by a diffuse-field compensation of the HRTFs, i.e., calculation of the DTFs. The DTFs were used in an auditory model for sagittal-plane sound localisation in order to predict the effect of the ear canal on sound-localisation performance.

In this investigation, we found strong spectral differences in HRTFs similar to [5], which are confirmed by measurements in [6]. We showed estimated sound-localisation errors as negligible when using DTFs. In the future, the effect of spectral differences in HRTFs will further be analysed by conducting an acoustic sound-localisation experiment.

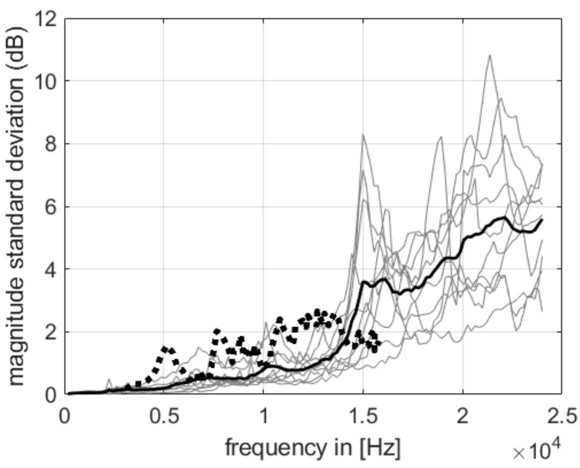


Figure 7: Standard deviations of the spectral differences (in dB) over the subjects. The bold dotted line represents the data from [2]. The grey lines represent our data, and the bold straight line their average.

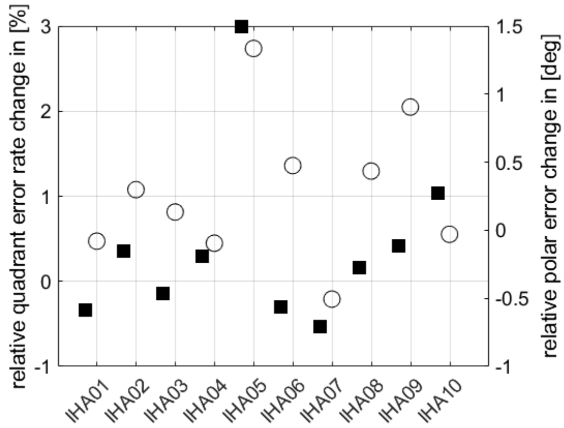


Figure 8: Differences in quadrant error rates (square, filled) and polar errors (circle, open) predicted for considered subjects using the DTFs of the eardrum as the template in the localisation model.

5. ACKNOWLEDGMENTS

This work was supported by the Austrian Research Promotion Agency (FFG, grant no. 871263, project "soft-pinna") and the European Union's Horizon 2020 research and innovation funding programme (grant no. 101017743, project "SONICOM").

6. REFERENCES

[1] J. Blauert, *Spatial hearing. The Psychophysics of Human Sound Localization*. Cambridge, MA: The MIT Press, revised edition ed., 1997.

[2] J. C. Middlebrooks, J. C. Makous, and D. M. Green, "Directional sensitivity of sound-pressure levels in the human ear canal," *The Journal of the Acoustical Society of America*, vol. 86, no. 1, pp. 89–108, 1989.

[3] D. Hammershøi and H. Møller, "Sound transmission to and within the human ear canal," *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 408–427, 1996.

[4] R. Roden and M. Blau, "The IHA database of human geometries including torso, head and complete outer ears for acoustic research," Sept. 2021. Funded by the Deutsche Forschungsgemeinschaft (DFG) – Project Nr. 352015383 – SFB 1330 HAPPAA C1.

[5] F. Di Giusto, D. Sinev, K. Pollack, S. van Ophem, and E. Deckers, "Analysis of impedance effects on head-related transfer functions of 3d printed pinna and ear canal replicas," in *Proceedings of Forum Acusticum*, (Turin), pp. 1–8, 2023.

[6] D. Sinev, F. Di Giusto, K. Pollack, K. Mick, and J. Peissig, "Assessment of the directional characteristics of the ear canal using 3d printed replicas and numerical simulations," in *Proceedings of Forum Acusticum*, (Turin), pp. 1–8, 2023.

[7] B. O. Community, *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018.

[8] T. Palm, S. Koch, F. Brinkmann, and M. Alexa, "Curvature-adaptive mesh grading for numerical approximation of head-related transfer functions," pp. 1111–1114, 08 2021.

[9] H. Ziegelwanger, W. Kreuzer, and P. Majdak, "Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions," in *Proceedings of the 22nd International Congress on Sound and Vibration*, (Florence, IT), pp. 1–8, July 2015.

[10] F. Brinkmann, W. Kreuzer, J. Thomsen, S. Dombrovskis, K. Pollack, S. Weinzierl, and P. Majdak, "Recent advances in an open software for numerical hrtf calculation," *Journal of the Audio Engineering Society*, 2023.

[11] P. Majdak, F. Zotter, F. Brinkmann, J. De Muynke, M. Mihocic, and M. Noisternig, "Spatially oriented format for acoustics 2.1: Introduction and recent advances," *J. Audio Eng. Soc*, vol. 70, no. 7/8, pp. 565–584, 2022.

[12] W. Kreuzer, K. Pollack, P. Majdak, and F. Brinkmann, "Mesh2hrtf / numcalc: An open-source project to calculatehrtfs and wave scattering in 3d," (Aalborg, Denmark), pp. 443–452, 2022.

[13] H. Ziegelwanger, P. Majdak, and W. Kreuzer, "Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization," *The Journal of the Acoustical Society of America*, vol. 138, pp. 208–222, July 2015.

[14] R. Baumgartner, P. Majdak, and B. Laback, "Modeling sound-source localization in sagittal planes for human listeners," *The Journal of the Acoustical Society of America*, vol. 136, no. 2, pp. 791–802, 2014.

[15] P. Majdak, C. Hollomey, and R. Baumgartner, "Amt 1. x: A toolbox for reproducible research in auditory modeling," *Acta Acustica*, vol. 6, p. 19, 2022.

[16] P. Minnaar, J. Plogsties, and F. Christensen, "Directional resolution of head-related transfer functions required in binaural synthesis," *J. Audio Eng. Soc*, vol. 53, no. 10, pp. 919–929, 2005.

[17] C. Jenny and C. Reuter, "Usability of individualized head-related transfer functions in virtual reality: Empirical study with perceptual attributes in sagittal plane sound localization," *JMIR Serious Games*, vol. 8, p. e17576, Sep 2020.