PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Variability of Head Related Transfer Functions across subjects

Maciej JASIŃSKI[1]; Jan ŻERA[2]

Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Radioelectronics and Multimedia Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland

ABSTRACT

The paper discusses across subject differences of measured Head Related Transfer Functions (HRTFs) in horizontal and vertical planes. The comparison was also done with reference to the results obtained with the use of an acoustic manikin. The work was done for an exemplary database available on-line and for the HRTFs measured with the use of a specific measurement setup in our laboratory. In the latter case, acquisition of HRTF pairs (left and right ear) was done at 370 sound source positions in the surrounding space with the use of two miniature microphones placed at the entrance to the blocked ear canal. Across subject differences of measured HRTFs were assessed in horizontal planes at various elevation angles. The results showed that HRTF is nearly symmetrical in the 300 Hz to 1 kHz frequency band. For frequencies above 1 kHz (more so above 5 kHz), as expected, the effect of interaural asymmetry increases due to filtering effect of complicated anatomical structures of pinnae. The influence of these effects differed among 15 subjects who participated in the study. The results obtained with the use of the acoustic manikin were not representative (as an average) for results with subjects.

Keywords: HRTF, binaural hearing

1. INTRODUCTION

Head Related Transfer Functions (HRTFs) contain all the information allowing the listener to precisely locate the sound source in the surrounding three-dimensional auditory space. As the HRTF features are person related and depend on the anthropometrical measures, the best auralization results are obtained with the use of individualized HRTF functions. When this requirement is not fulfilled sound localization errors increase and the quality of the spatial auditory image worsens (1,2). The most reliable method of obtaining personalized HRTF is by direct measurement, which is, in general, time consuming and requires specialized measurement setup. Therefore, other approaches try to acquire usable HRTFs by either using an acoustic manikin, by modeling, or by selecting the most appropriate function from existing databases, e.g. (3–5).

A general need for personalization of the HRTFs with the use of relatively simple methods has increased interest in assessing the differences among various HRTFs and finding general measures of the variability of HRTFs in horizontal and vertical planes (6–9). To assess the variability of the HRTF magnitude Jo et al. (6) tested the CIPIC database using spectral measure equivalent to Euclidian distance metrics, similar to the measure described further in this paper. Wideband analysis and analysis in 1/3-octave bands showed that large variability among the HRTFs occurs in front and back sound positions while small variations occur in high elevation angles. Zhong et al. (7) assessed HRTF variability for a group of 52 subjects measured in the laboratory, applying spectral analysis in the ERB bands (10). A band related interaural cross-correlation function was used to measure the symmetry of HRTFs in horizontal and median planes. In left-right direction, in condition in which the sound source was on the same side as the ear measured, the symmetry was greater than in condition in which the sound source was on the opposite side to the ear measured. In the former condition the HRTFs were symmetric in a frequency range up to 5 kHz, whereas in the latter condition only to 1 kHz. In median plane, a noticeable asymmetry occurs above 6 kHz. Genovese et al. (8,9) calculated the interaural time difference from the HRIRs (Head Related Impulse Responses, corresponding to HRTFs) of four major HRTF databases, assessed their variability in horizontal and vertical planes, and found largest differences of about 90 μs at 110° horizontally, and the differences decreased with elevation angle.

[1] M.Jasinski@ire.pw.edu.pl
[2] J.Zera@ire.pw.edu.pl

The aim of this study was to assess the variability of HRTFs measured on a group of subjects, for the data taken from the LISTEN HRTF database (4), and with an acoustic manikin (HATS, Head and Torso Simulator). The comparison of HRTFs was done along horizontal (azimuth) and vertical (elevation) angles.

2. The HRTF analysis

In total, 66 HRTF pairs were analyzed. This included 15 subjects participating in our own measurements, the 50 sets of data from the LISTEN database (4), and HRTFs obtained with the use of Head and Torso Simulator (HATS, acoustic manikin).

2.1 Measurements of the HRTF

The HRTF measurements were conducted using similar measurement setup to that used in a previous study (11). The sound source was in fixed position and the subject was rotated in the sound field. The frequency responses were recorded with miniature microphones placed in blocked ear canals.

The measurement setup consisted of an arc with 10 wideband (300 Hz−15 kHz) loudspeakers (GD 12/8/2), fixed in position. The subjects and the acoustic manikin (Brüel & Kjær, B&K 4100-D) were rotated, with the use of the turntable (B&K 3921), against the loudspeakers placed at a distance of 1.5 m measured from the center of the subject's or the manikin's head.

The MLS signals were recorded using a pair of B&K 4101-A miniature binaural microphones. The microphones, 5.4 mm in diameter, were fixed at the entrances to the subjects' ear canals. In the case of the B&K 4100-D manikin, HRTFs were recorded with B&K 4189 1/2-inch microphones which are the internal manikin microphones placed in a position corresponding to the entrance of the ear canal.

2.2 Subjects

Fifteen subjects, four women and 11 men, aged 19-29 years, participated voluntarily in the study. None of them had any visible asymmetry in the pinnae shape or their positions on the head. The measurements with subjects (and the manikin) were made in 370 sound source locations, at 36 azimuth angles (from 5° to 355° in 10° steps plus 0° angle) and 10 elevation angles from -45° to 90° in 15° steps.

2.3 LISTEN Database

LISTEN Database contains a set of HRIRs (Head Related Impulse Responses) of 51 subjects. For each subject, the database includes 187 HRIR pairs measured with different horizontal density in 10 evenly spaced elevation angles ranging from -45° to 90°, in 15° steps. In the horizontal plane, the data covered a range from 0° to 360° in 15°, 30°, and 60° steps, respectively for elevation angles from -45° to 45°, at 60°, and 75°. A single HRIR pair represents a 90°elevation angle (4).

The LISTEN database was chosen as its data structure represents similar angles in the vertical plane. Only 50 subjects were taken from the database as the data for one of the subjects are flawed by artifacts, which was noted by Genovese et al. (8).

2.4 Root Mean Square Difference

The comparison of the HRTFs was made by calculating the Root Mean Square Difference (RMSD) of a HRTF pair, according to the equation (1):

$$RMSD = \frac{1}{\sqrt{n}} \sqrt{\sum_n (HRTF_{L,n} - HRTF_{R,n})^2}, \quad (1)$$

where $HRTF_{L,n}$ and $HRTF_{R,n}$ are the HRTFs for the left and right ears, and n is the number of discrete frequency points in the HRTF vector. This measure was also used in our previous study (11) and is similar to that used by Jo et al. (6).

3. Results

3.1 Horizontal plane

Figure 1 shows the distribution of the RMSD values in a semi azimuthal plane, at 0° elevation angle. The upper row represents a condition in which the ear is on the same side as the sound source, i.e. the source is on the left side for the left ear, and on the right side for the right ear (both ears directly stimulated by the sound). The lower row represents a condition in which the ear is on the opposite side of the sound source, i.e. for the left ear the source is on the right side, and for the right ear the source is on the left side (ears shadowed from the sound).

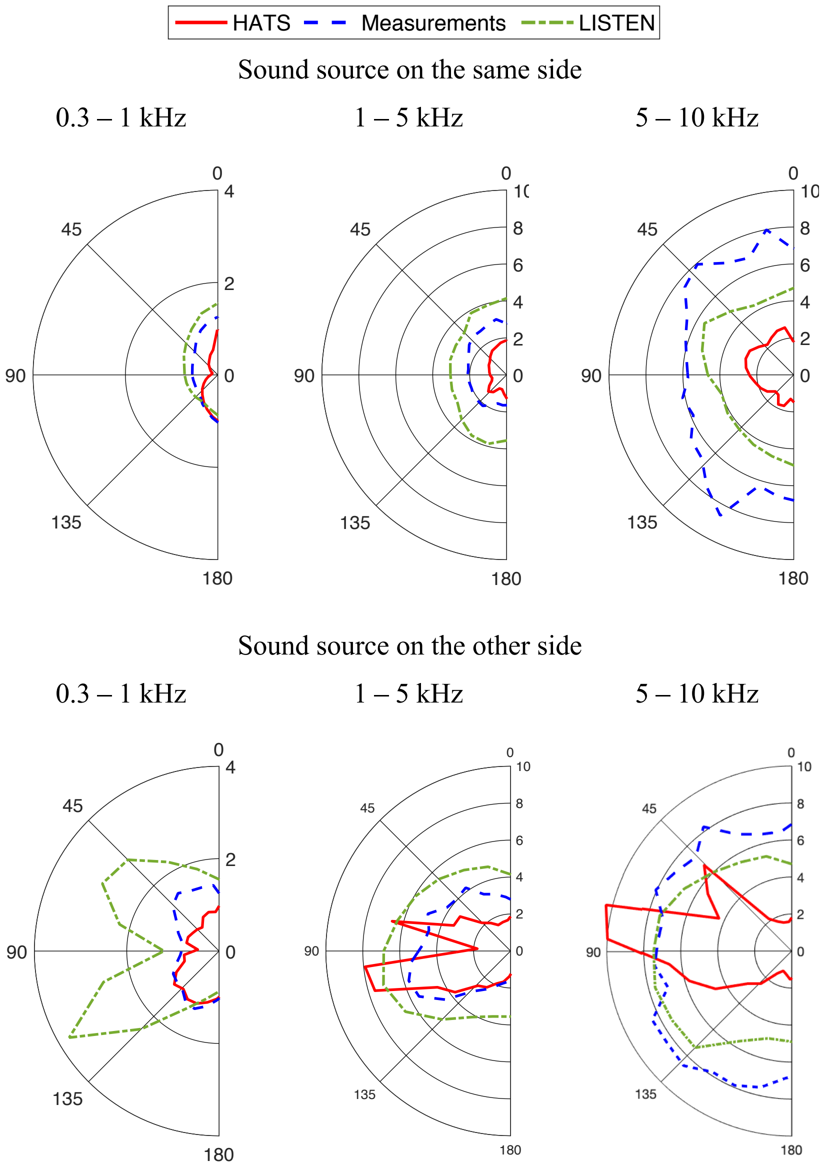

Figure 1 – RMSD measure of the left-right ear difference as a function of the semi azimuthal angle at 0° elevation. Upper row: the ear on the same side as the sound source. Lower row: the ear on the opposite side of the source (both ears shadowed). Conditions: mean over measured subjects (blue dashed line), mean over LISTEN database (green dash-dot line), acoustic manikin (red solid line)

The largest RMSD values were observed in the highest frequency range of 5-10 kHz (Figure 1, lower left panel). For the HRTFs measured in our laboratory (blue dashed line) the RMSD is 6-8 dB. For the data from the LISTEN database (green dashed line) it is by about 1-2 dB smaller than for measured HRTFs, in front and back direction (0° and 180°). Otherwise, the RMSD for measured and database data are similar. An important difference is apparent for the manikin (red solid line) for which, at side angles, the RMSD is as large as 10 dB and for front and back angles is less than 2 dB. The large RMSD for side angles must be related to a difference in the angle of the manikin left and right side position against the sound source during the measurements, which is not that important when the source is placed in the median plane (0° or 180°). At this angle, the simplified artificial manikin's symmetry is dominating. For human subjects, the asymmetrical geometry of head and face does not allow for such a decrease in RMSD in the 5-10 kHz range (wavelength 3-6 cm) in front-back position of the sound source. Therefore, the individual RMSDs of subjects measured and taken from the database differed from each other and from the manikin's RMSD, which resulted in a more even distribution in semicircle shown in Figure 1. Among subjects measured not more than five people showed the RMSD curves similar to that of the manikin's (average difference within 2-3 dB), and for four subjects the differences were much larger (average difference 6-9 dB).

In mid frequency range of 1- 5 kHz shown in Figure 1, middle lower panel, the conditions are qualitatively similar except that the RMSD values are smaller (3-4, 4-7, and 2-8 dB for the measured subjects, the database, and manikin, respectively). The main difference is that RMSD values in front-back position of the sound source decreased more than for the side position, as compared to the 5-10 kHz frequency range (averaged subjects' results became more like the manikin's data). The possible explanation of this finding is that in lower frequency range, thus for wavelengths of 6-30 cm, the individual face and head asymmetry among people are not that important.

In mid frequency range of 0.3-1 kHz shown in the left lower panel in Figure 1, the RMSD values are much smaller than for the two higher frequency ranges. For the manikin (red solid line), and subjects measured in the laboratory (blue dashed line), the RMSD is less than 1.5 dB. For the LISTEN database the RMSD ranges from the 1.5 to 4 dB. For all conditions, there is a decrease in RMSD (much stronger for the database) at the 90° position of the sound source, which is likely a result of dominant role of the direct wave reaching the entrance of the ear canal with minimum interference from the pinnae, due to large wavelength.

In the condition of the both ears exposed to the sound source shown in upper panels in Figure 1, the RMSD values are less than 1.5 dB in the low frequency range (upper left panel) and less than 4 dB in the mid frequency range (upper middle panel). Some decrease in the RMSD is also seen in a frequency range of 5-10 kHz (upper left panel in Figure 1). Unlike the conditions shown in lower panels of Figure 1, the pattern of the RMSD angular distribution is similar for subjects (either measured or for HRTFs from the LISTEN database) and the manikin. For all frequencies the RMSD for the manikin is less than 2 dB and is much smaller than for subjects. This indicates an importance of indirect waveform reaching the side of the head for creating the differences in the measured HRTFs, and again the significance of the geometrical asymmetries among heads and pinnae (as for perfectly symmetric manikin the RMSD values in "exposed" condition are definitely small).

3.2 Vertical planes

Figures 2 and 3 show the distribution of the RMSD values for all elevation angles, respectively in median plane (azimuthal angle 0°-180°) and transverse plane (azimuthal angle 90°-270°). In Figure 2, the upper row represents conditions in which the sound source is at the front of the head, and the lower row represents conditions in which the ear is behind the head. In Figure 3, the conditions are similar to those in Figure 1, the upper row represents conditions in which the ear is on the same side as the sound source, and the lower row represents conditions in which the ear is on the opposite side to the sound source. As a general tendency, in data for both vertical planes, it may be noticed that in horizontal plane with a 0° elevation angle the largest RMSD values were obtained in high frequency range of 5-10 kHz. In addition, the lowest value was obtained for the 90° elevation angle, i.e. with the sound source positioned over the head.

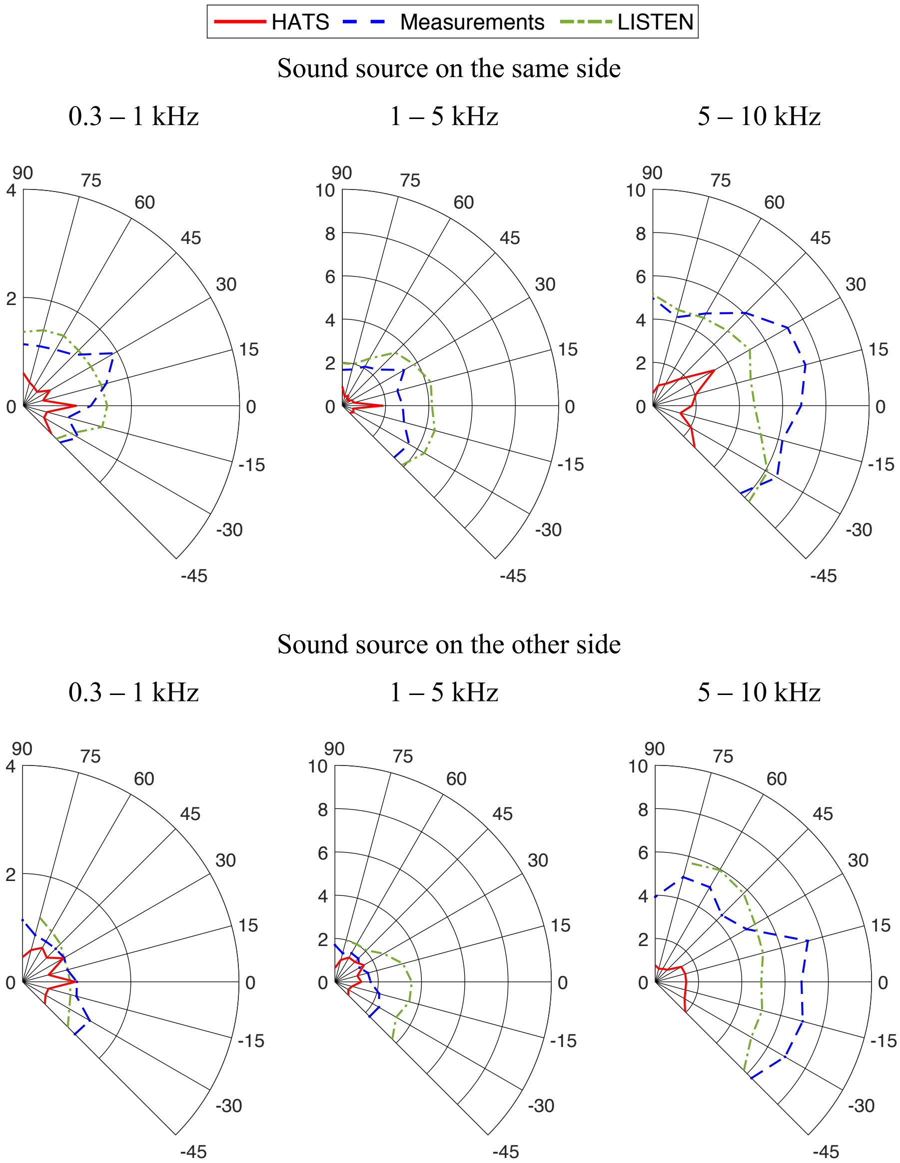

Figure 2 – RMSD measure of the left-right ear difference as a function of the elevation angle in median plane. Upper row: the sound source in the frontal position (0° azimuth). Lower row: the sound source in the back position (180° azimuth). Conditions: mean over measured subjects (blue dashed line), mean over LISTEN database (green dash dot line), acoustic manikin (red solid line)

In the frontal sound source position (Figure 2, upper panels), in each frequency range the RMSD for people (either measured subjects or data from the database) is much larger than for the manikin. For 5 – 10-kHz frequency range, the largest RMSD for subjects measured in the laboratory occurs at an elevation angle of 15° - 30° above the horizontal plane, whereas for the database at -30° (below the horizontal plane). For mid and low frequencies (Figure 2, upper middle and left panels) there is a similar difference, but less pronounced. In the back sound source position (Figure 2, lower panels) and for the 5-10 kHz range (left panel) the RMSD is larger for negative elevation angles for HRTFs measured in the laboratory, and larger for positive elevation angles for data from the database. In mid frequencies (Figure 2, lower middle panel) the RMSD is less than 4 dB for data from the database, less than 2.5 dB for subjects measured in the laboratory and especially small (below 1.5 dB) for the manikin. It is apparent that, in the low frequency range (Figure 2, lower lest panel), the RMSD is less than 1.5 dB for subjects and less than 1 dB for the manikin. In summary, the data seen in the lower panel in Figure 2 show that the localization of the sound source in the back head position may be ineffective. But on the other hand, the variability in HRTF functions is the smallest being not critical for HRTF personalization.

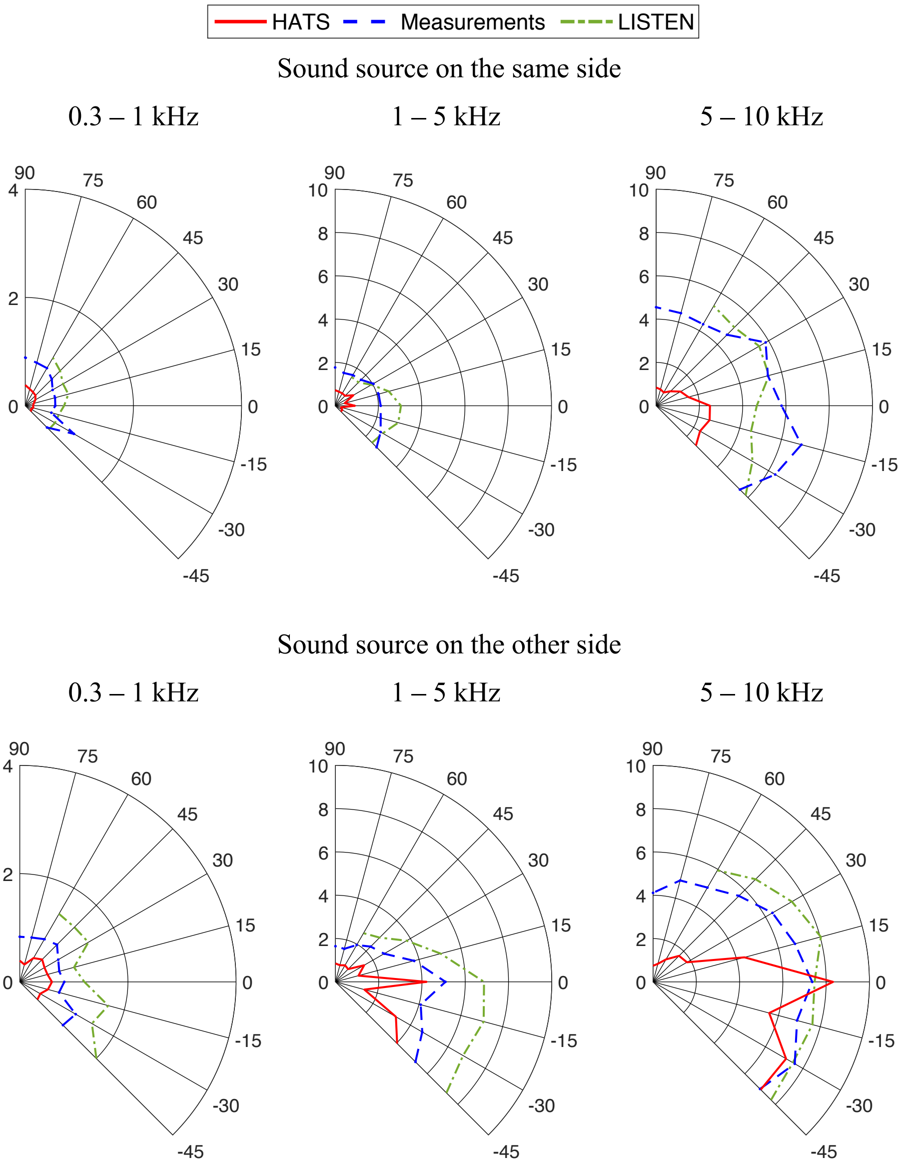

Figure 3 – RMSD measure of the left-right ear difference as a function of the elevation angle in the transverse plane (90°-270° azimuth). Upper row: the ear on the same side as the sound source. Lower row: the ear on the opposite side to the source (both ears shadowed). Conditions: mean over measured subjects (blue dashed line), mean over LISTEN database (green dash dot line), acoustic manikin (red solid line)

Left-right conditions presented in Figure 3 are essentially quite different from those shown in Figure 2. In Figure 3, the 0° elevation angle corresponds to 90° horizontal angle in Figure 1. In Figure 2, the 0° elevation angle corresponded to either 0° or 180° horizontal angle in Figure 1 (either upper or lower panels). Therefore, unlike as in the Figure 2 data for the manikin at 5-10 kHz (Figure 3 lower right panel) show large (more than 8 dB) RMSD for 0° of elevation angle. This value quickly drops to less than 1 dB for elevation angle increased to 90°. In summary, data shown in Figure 3 for the left-right position of the sound source similarly to that front back source position in Figure 2, indicate that variability in HRTF functions vastly decreases for elevated position of the sound source up to the position above the head (90°).

4. CONCLUSIONS

In this study, the variability of HRTFs measured on a group of subjects, on an acoustic manikin (HATS, Head and Torso Simulator), and the HRTFs obtained from the LISTEN database were assessed. Results showed that the variability among subjects measured by the proposed RMSD parameter is large in a frequency range above 5 kHz, slightly decreases in 1-5 kHz range and is minimal below 1 kHz. The results of measurements in our laboratory were not significantly different from the data of LISTEN database.

The study showed that the manikin does not represent mean spectral RMSD values of the human subjects' population. Thus, even in such a statistical sense, the manikin's HRTFs cannot be successfully used for creating spatial images. Only in specific circumstances one may expect a similarity of the manikin's and a subject's HRTFs.

Symmetrical heads, like that of the HATS' are characterized with large geometrical symmetry across the median plane. This results in small RMSD values with an increase of the elevation angle, and thus small distinction, both in horizontal and vertical planes, of sound sources located as elevated over the horizontal plane. The right-left ear symmetry, across all frequency bands, also increases for subjects with the increase of the elevation angle of the sound source, however to lesser extent than for the manikin.

The results show that HRTF might be considered left-right ear symmetrical in the lower frequency band (300 Hz – 1 kHz). The symmetry is likely compensated by the Interaural Time Difference (ITD) in this frequency range.

ACKNOWLEDGEMENTS

This work was supported by the statutory grant 504/04064/1034/40.00 from the Warsaw University of Technology.

The first author's participation in the ICA 2019 congress was supported by:
"Project PROM - International Scholarship Exchange for PhD Students and Academic Staff" is financed from the European Social Fund under the Operational Program Knowledge Education Development, non-competitive project entitled International Scholarship Exchange for PhD Students and Academic Staff, contract number: POWR.03.03.00-00-PN13/18.

European Funds Knowledge Education Development | Republic of Poland | NAWA Polish National Agency for Academic Exchange | European Union European Social Fund

REFERENCES

1. Geronazzo M, Spagnol S, Avanzini F. Do we need individual head-related transfer functions for vertical localization? the case study of a spectral notch distance metric. IEEE/ACM Trans Audio Speech Lang Process. 2018;26(7):1243–56.
2. Wenzel EM, Arruda M, Kistler DJ, Wightman FL. Localization Using Non-individualized Head Related Transfer Functions. J Acoust Soc Amer. 1993;94:111–23.
3. Algazi VR, Duda RO, Thompson DM, Avendano C. The CIPIC HRTF database. In: Proceedings of the IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics. 2001. p. 99–102.
4. Warusfel O. LISTEN HRTF Database [Internet]. 2003. Available from: http://recherche.ircam.fr/equipes/salles/listen/
5. ARI HRTF Database [Internet]. Available from: https://www.kfs.oeaw.ac.at/index.php?view=article&id=608&lang=en
6. Jo H, Park Y, Park Y. Analysis of individual differences in Head-Related Transfer Functions by spectral distortion. 2009 ICCAS-SICE. 2009;1769–72.
7. Zhong X, Zhang F, Xie B. On the spatial symmetry of head-related transfer functions. Appl Acoust. 2013;74(6):856–64.
8. Genovese AF, Juras J, Miller C, Roginska A. Investigation of ITD symmetry in measured HRIRS. In: The 22nd International Conference on Auditory Display (ICAD–2016). Canberra, Australia; 2016.
9. Genovese AF, Juras J, Miller C, Roginska A. The Effect of Elevation on ITD Symmetry. 2016;
10. Moore BCJ. An introduction to the psychology of hearing. 5 th. edit. San Diego, USA: Academic Press; 2003.
11. Jasiński M, Żera J. Inter Subject Differences of Head Related Transfer Functions. In: 2018 Joint Conference - Acoustics. Ustka, Poland; 2018.