Variance in Measured Binaural Room Transfer Functions of Individuals

Stephan Werner[1], Martina Böhme[1], Thomas Mayenfels[1], and Florian Klein[1]

[1] *Electronic Media Technology Group, Ilmenau University of Technology, Ilmenau, Germany*
E-mail: stephan.werner@tu-ilmenau.de

Introduction

Binaural synthesis systems allow the reconstruction of an acoustic environment by employing binaural room transfer functions (BRTFs) and convolving them with dry audio signals. The aim of this technique is to create a plausible and immersive auditory illusion with a high quality of experience [1]. The perceived quality features such as localization, externalization, or coloration should be comparable to those of a real listening situation [2, 3]. The experienced quality depends, amongst others, on the personalization of the binaural resynthesis [4]. The smaller the deviation of the used BRTF for the synthesis from the individual transfer function, the better the perceived quality. Recorded BRTFs have been shown to exhibit high individual differences [5].

This contribution presents an analysis of individually measured BRTFs for several subjects, rooms, and source positions. A dataset of 25 BRTFs was used to investigate for what frequency ranges particularly large individual differences or slight differences occur. The used dataset is stored in the SOFA (Spatially Oriented Format for Acoustics, [6]) file format and is available at http://www.tu-ilmenau.de/mt-emt/forschung/andere-projekte/.

Binaural Recordings

For the creation of the dataset, the individual BRTFs of 25 subjects were recorded in two rooms with different acoustic properties for six sound source positions. The recordings were taken using the microphones from the Smyth Realiser system [7] at the blocked ear canal. During the measurement, no head movements of the test person were allowed.

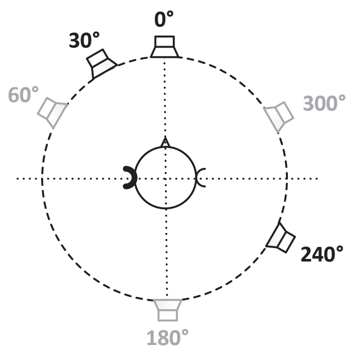

Figure 1 – Sound source positions and recording position used for measurements; radius approx. 2.2m; only the BRTFs for 0°, 30°, 240° measured at the left ear are analyzed in this paper.

Figure 1 shows the positions of the sound sources and the recording position in the horizontal plane. As sound sources, six Genelec 1030A loudspeakers were circularly arranged around the recording position with a radius of 2.2m. The following sound source directions were chosen: 0°, 30°, 60°, 180°, 240°, and 300°. These directions enable investigations on the perceived directional confusions in the frontal and back quadrants. For the sake of simplicity, only the analysis of the BRIRs measured at the left ear of the test subjects for the directions at 0°, 30°, and 240° are presented in this paper.

The recordings of the BRIRs were conducted in a listening lab compliant to ITU-BS.1116-1 (HL; $V = 179\ m^3$; $T60 = 0.34$ s) and a depleted seminar room (SR; $V = 182\ m^3$, $T60 = 2.0$ s). The Energy Decay Curves (EDCs) for both rooms are depicted in Figure 2. The EDCs were measured using an omni-directional microphone and one loudspeaker (Genelec 1030A) at the 0°-position. The measurements do not fulfill the requirements of a standardized room acoustic measurement since directional loudspeaker at only one position in the room was used. Therefore, the EDCs shown in Figure 2 are steeper than usual for a standardized measurement. However, this has no impact on the analysis of the BRTFs in this paper.

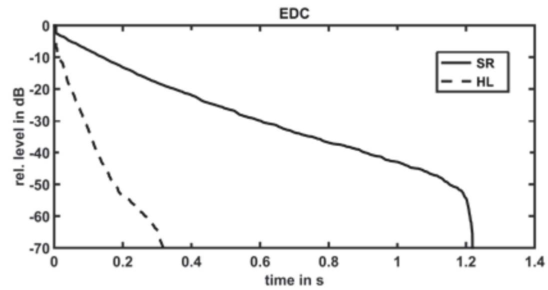

Figure 2 – Energy Decay Curves (EDCs) for the recording rooms listening lab (HL) and seminar room (SR).

Measurement Results

In Figure 3, the individually measured BRTFs at the left ear of the subjects are comparatively displayed for both the listening lab and the seminar room. The results are presented in graphs showing the relative amplitude at the blocked ear canal over the frequency. The results for all 25 subjects are displayed as quantiles. As stated before, the shown results focus only on the left ear. Commonly, the human being's anatomy is not perfectly symmetrical which leads to differences in the frequency responses of the left and right ear. As a qualitative representation, this fact is not taken into account here for reasons of simplicity, but should always kept in mind.

On the left side of Figure 3, the quartiles as well as the 0.95- and the 0.05-quantile for the measurements of the listening lab are shown for the directions 0°, 30°, and 240°. Thereby, the 0.5-quartile, or the median, is the central value of the data distribution, i.e., there is an equal number of values above and below the median. On the right side, the same quantiles for the measurements of the seminar room can be seen.

Listening lab (HL) **Seminar room (SR)**

0° 30° 240°

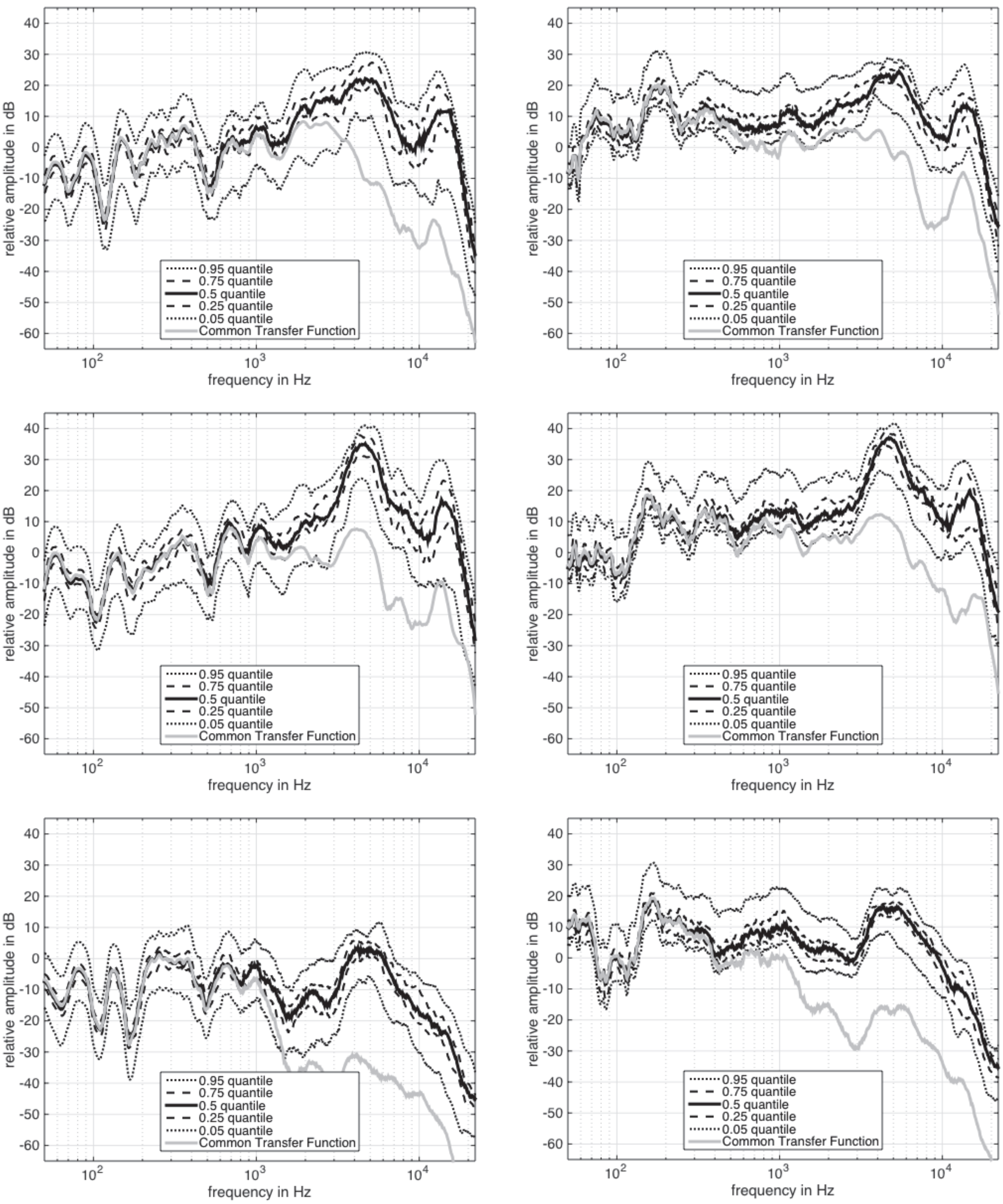

Figure 3 – Measured BRTFs at the left ear of all subjects as quantiles for the directions 0°, 30°, and 240° for the listening lab on the left side and the seminar room on the right side, respectively.

As can be seen from the difference between the 0.95- and 0.05-quantile, the BRTFs exhibit differences of about 20 dB over the entire frequency range. Within low and mid frequencies, the relative amplitude is lower in the listening lab compared to the seminar room, which can be attributed to the dry room acoustics of the listening lab. For high frequencies, the BRTF curves of the same directions for both rooms are about the same.

Additionally, the Common Transfer Function (CTF) was computed as the mean BRTF for each direction and room averaged over all test subjects.

It can be seen that up to about 0.8 kHz the median and the CTF correspond approximately. Within this frequency range, there are only minor individual differences. However, with higher frequencies the median and the CTF diverge. This may be explained by the fact that shadowing occurs when a quarter of the sound wavelength is shorter than the head size.

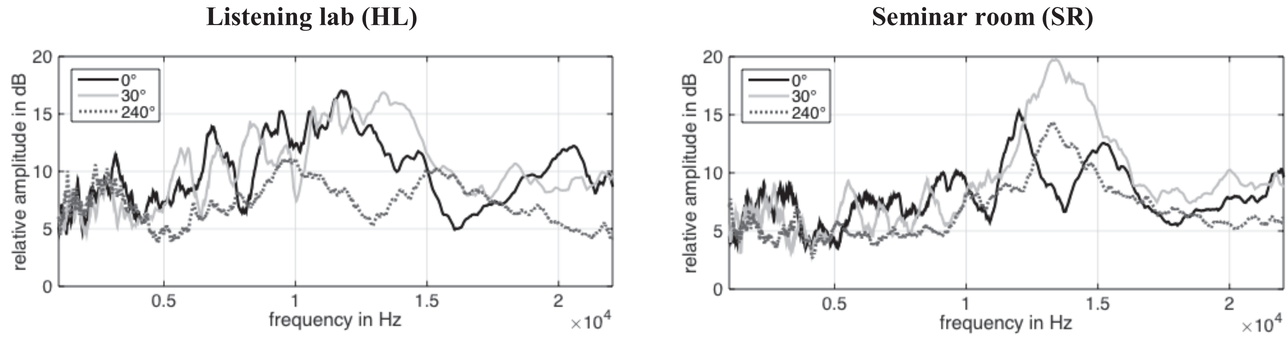

Figure 4 – Interquartile Distances (IQD) of the measured BRTFs of the left ear of all subjects for the directions 0°, 30°, and 240° for the listening lab (left) and the seminar room (right)

Due to the individual shape of the head and the pinnae, the individual differences in the BRTFs increase for frequencies higher than about 0.8 kHz. Large individual differences have a stronger impact on the mean BRTF than on the median BRTF, which is more stable concerning outliers. For this reason, the CTF deviates from the median BRTF for higher frequencies.

To illustrate the individual differences in more detail, the Inter-Quartile Distances (IQDs) for frequencies above 1 kHz, i.e., the frequency range where median and CTF diverge, were computed as well. The IQD is the difference between the 0.75-quantile and the 0.25-quantile. The IQDs for each direction and room are depicted in Figure 4. It allows some comfortable observation of the overall behavior and the differences in the frequency responses. The listening lab with its short reverberant time provides a slightly exponential increase to a maximum level difference of approximately 12 dB at around 10 kHz on 30°. Also to be mentioned are several dips of about 10 dB in this range at 0° and 30° as well. On 240°, the aforementioned maximum splits up into two lower maxima at about 1 kHz and 1.5 kHz. Taken as a whole, the frequency response for 240° is smoother than those of the other directions. In contrast, the reverberant room shows a nearly constant behavior except a sharp peak of 15 dB at around 10 kHz, also 30°. At 240°, a similar but about 5 dB lower peak is recognizable. In this room, the splitting into two peaks is happening at 0° in a slightly sharper manner. By comparison, a smoothing of individual differences over the frequency range by a longer reverberation time seems to happen. The presence of peaks at the particular frequencies points out remarkable resemblances in the individual anatomy of the pinnae supported by the distinct room acoustics.

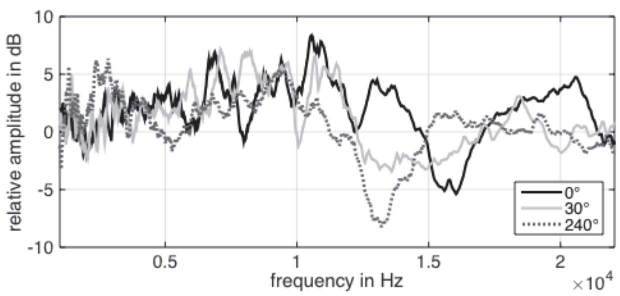

Figure 5 –Difference between the IQDs of the listening lab and the IQDs of the seminar room.

In Figure 5, the differences between the IQDs of the listening lab and the seminar room are computed, respectively. In this representation, the particular room characteristics primarily evoked by the reverberation time should become clearer. As mentioned before, the main differences occur in a frequency range from 10 kHz to 17 kHz, which strongly vary within the source directions.

Conclusions

In first place, the exemplary depiction of averaged BRTFs gives a quite informative overview on general frequency dependent level differences in human hearing and so does provide additional confirmation on the findings of Blauert [8] and many other researchers who focused on localization and its dependency on certain audible frequencies.

When going more into detail, the quite large differences in individuals will become visible. Based on the frequency dependent level variations of averagely 10 dB, the importance of individual BRTFs and their influence on the overall quality of experience becomes more obvious.

References

[1] Nicol, R., "Binaural Technology", Monographs of the American Engineering Society, AES 2010.

[2] Lindau, A., Brinkmann, F., Erbes, V., Lepa, S., Maempel, H.-J., Weinzierl, S., "A focus group approach towards a Spatial Audio Quality Inventory for virtual acoustic environments", EAA Joint Auralization and Ambisonics Symposium, Berlin, 2014.

[3] Le Bagousse, S., Colomes, C., Paquier, M., "State of the Art on Subjective Assessment of Spatial Sound Quality", Audio Engineering Society Conference, 38th Int. Conf.: Sound Quality Evaluation, 2010.

[4] Werner, S. and Klein, F.: "Influence of Context Dependent Quality Parameters on the Perception of Externalization and Direction of an Auditory Event", 55th AES Conf. Spatial Audio, Helsinki, Finland, 2014.

[5] Møller, H., Sørensen, M. F., Jensen, C. B., and Hammershøi, D.: "Binaural technique: Do we need individual recordings?", J. Audio Eng. Soc, 44, pp. 451-469, 1996.

[6] Audio Engineering Society Standard: AES69-2015 AES standard for file exchange - Spatial acoustic data file format, 2015.

[7] Smyth Research, HTM-1 miniature microphones: http://www.smyth-research.com, 29.03.2016.

[8] Blauert, J.: "Spatial Hearing- Rev. Ed.: The Psychophysics of Human Sound Localization", MIT Press, 1996.