INTER-NOISE 2006

3-6 DECEMBER 2006
HONOLULU, HAWAII, USA

Spatial and spectral properties of the dummy-head during measurements in the head-shadow area based on HRTF evaluation

György Wersényi[a]
Széchenyi István University
Department of Telecommunications
H-9024, Gyor
Egyetem tér 1.
Hungary

ABSTRACT

In accurate and repeatable measurements dummy-heads are often used to model the average human head and body. They are suited for standardized measurements and for investigating the human spatial hearing and localization performance. The monaural Head-Related Transfer Functions (HRTFs) of the dummy-head can be used for various investigations. This paper uses the HRTF-set of a Brüel & Kjaer head and torso simulator focusing on the so called monaural head-shadow area, where one of the ears is shadowed by the head itself. Based on long-term measurements using the bare torso as well as other accessories (glasses, clothing etc.) on it, the extent of the head-shadow area will be presented in frequency and space. The head-shadow area is investigated in connection with the overall SNR of the measurement and sensitivity domains of the ears. Conclusions are drawn for binaural recognition in human spatial hearing using low-frequency 'bright spots' and high-frequency information during lateral-contralateral evaluation.

1 INTRODUCTION

The Head-Related Transfer Functions (HRTFs) describe the transmission from a given point in the free-field to the eardrums. This filtering effect is responsible for basic localization cues during human spatial hearing [1-15]. The measurement and reproduction of these transfer functions are elementary problems in hearing research. Measurements can be done on real human heads by placing small microphones on the eardrum or at the entrance of the ear-canal. Such measurements deliver individual results but human subjects are not very well suited for long-term measurements. Furthermore, repeatability and reproduction of the results are hard to realize.

Instead of real humans dummy-heads (also known as head and torso simulators) are often used in measurements. They model the average human head and body as well as its acoustical properties (such as ear canal length, eardrum impedance etc.). A measurement system equipped with a dummy-head is suited for long-term acoustical and noise measurements having the advantageous property of standardized and repeatable measurements.

After recording a set of HRTFs, they will be implemented in a playback system using equalized headphones and real-time HRTF filtering (in the time domain or in the frequency domain). Listeners report about sensation, direction of simulated sound sources and localization blur. In binaural listening tests it is common and well-known that dummy-head HRTFs are inferior to individual HRTFs, thus, localization results tend to be better using individually recorded HRTFs. Furthermore, there are other localization cues to solve localization tasks and problems like in-the-head localization that is also present during headphone playback.

---

[a] Email address: wersenyi@sparc.core.hu

For investigating the role of the HRTFs in the human spatial hearing an exhaustive investigation was made using a dummy-head for recording the HRTFs in one-degree spatial resolution horizontally and in 5 degrees resolution vertically [16]. Results were already presented about the effect of hair, clothing and the environment near the head [17-19]. These results suggest improvement of the playback system and reveal the important role of the cortex and higher processing in the brain during solving localization tasks.

This paper analyzes the monaural (left ear condition) HRTFs of a dummy-head focusing on the repeatability property in frequency and space. The evaluation of directional information and the role of the head-shadow area are discussed. Finally, some words and conclusions are drawn for binaural playback systems.

2 MEASUREMENT SETUP

The measurement setup includes the Brüel & Kjaer Head and Torso Simulator Type 4128 placed on a turntable in the anechoic room. The turntable is controlled by a computer in 1 degree steps. Accuracy and repeatability was deeply investigated in order to create a measurement system suited for long-term accurate measurements [20]. Changes of 1 dB in the measured transfer functions can be evaluated. Pseudo random white noise signal was used as stimuli and results were collected for both ears simultaneously.

The measured data was accumulated and averaged and after applying the FFT the magnitude of the transfer functions were plotted as function of frequency. The HRTFs were calculated as usual:

$$HRTF = \frac{P_1(jw)}{P_2(jw)} \tag{1}$$

where P_1 is the sound pressure at the eardrum and P_2 is the sound pressure in the origin of the head-related coordinate system at the same signal and sound source, but recorded with a unidirectional microphone [2].

The analysis uses the following definition of the free-field HRTF Difference (HRTFD). It is defined as a quotient of HRTFs from the same direction but under modified conditions:

$$HRTFD = \frac{HRTF_{C1}}{HRTF_{C2}} \tag{2}$$

where C_2 identifies the reference and C_1 the modified condition. We plot the 20log/HRTFD/ magnitude response as the function of frequency or as 2D polar histogram as function of frequency and azimuth [17-19]. With simple words: an HRTFD is the difference between re-measured HRTFs from the same direction. They are well suited for investigating the repeatable property of the measurement system as well.

The complex quotient refers to subtraction of two logarithmic magnitude responses. This difference gives us the deviation in dB between two HRTFs. For analyzing the HRTFDs we do not need individual recordings on real human heads because the dividing will eliminate the individual differences. Due to the symmetry of the dummy-head, only results for one ear will be presented.

3 RESULTS

3.1 The Sensitivity Domains

The companion paper [21] showed the monaural HRTFs of the dummy-head plotted in 1 degree resolutions. We could conclude that

- spatial domains, where the overall signal level of the HRTFs is high, is where the localization blur is small. This can be simply explained by the high SNR [22],
- domains, where HRTFs can be re-measured with high accuracy the spatial separation capability (sensitivity) of the ears is good,
- the head-shadow area decreases the localization performance because high frequency evaluation (above about 1600 Hz) is not possible.

These results are based on high accuracy HRTF measurements using the dummy-head. It has to be mentioned that binaural evaluation is made by two ears simultaneously.

An engineer can handle the ears as ordinary antennas. Antennas have directional filtering, with other words, directional dependent sensitivity. The most important is the pinna, followed by the size and geometry of the head and body. Fig.1 shows the role of the pinna filtering effect at frontal incidence. Two HRTFs were measured with and without the artificial pinna of the head and torso simulator. The sound collecting effect at 3 kHz and above 8 kHz is significant. An average differences between the spectra of the torso below 3 kHz with and without pinna of 0,86 dB was reported in [23]. Our measurement could not show differences less than 0,5 dB.

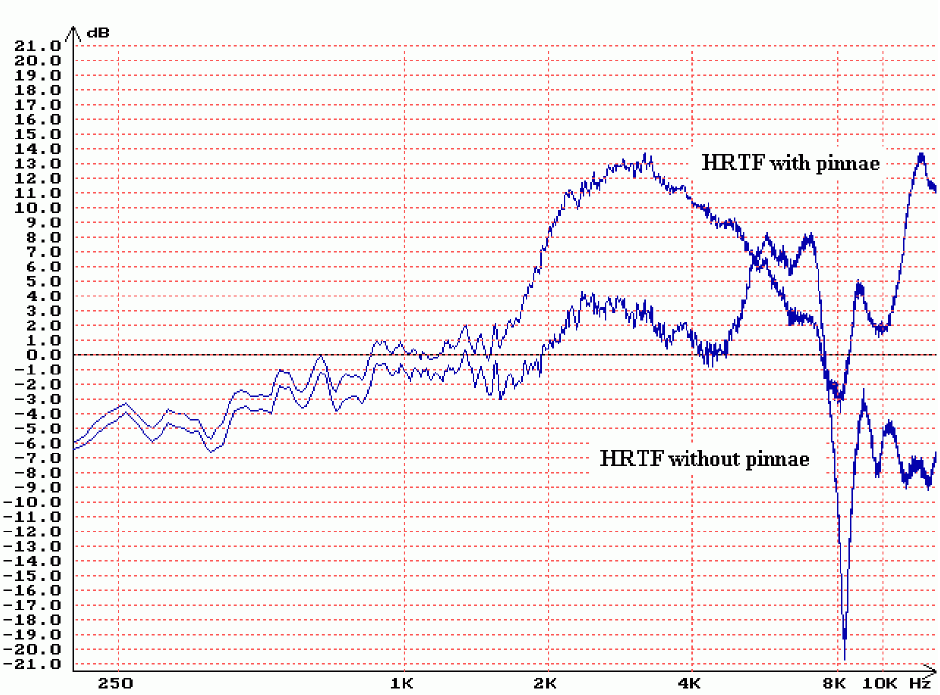

Fig.1. Effect of the pinna at frontal incidence. Both HRTFs contain the effects of the torso and the head. The reflecting and amplifying effect of the pinna is clearly visible at the main resonance frequencies of 3, 9 and 11 kHz.

The most sensitive spatial area is in the frontal direction ±20°. In this domain both ears' HRTFs are very helpful for evaluating directional information in the entire frequency range. This seams to be smaller than the stereo-area of the eyes. Based on the overall signal level (the highest

SNR) provided by the HRTFs, the direction 45° seems to be the most sensitive direction. This has to be in connection with the placement of the pinna on the head (see Fig.2).

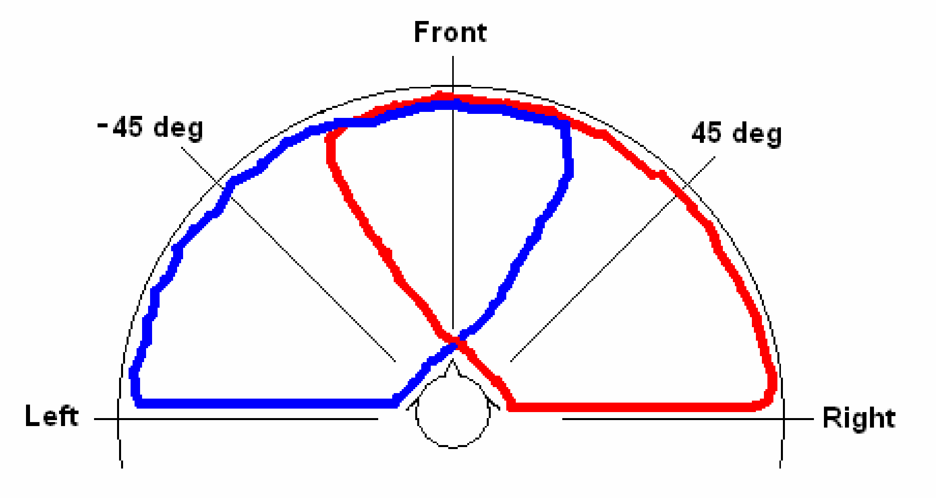

Fig.2. Footprints of the ears representing the monaural sensitivity regions based on the HRTF analysis of the dummy-head.

Our measurement shows that this kind of "monaural sensitivity region" can be recognized significantly only between elevations of -10° to +30°. This seems to be logical, as above +30° the shadowing effect of the head disappears.

Local increase of the localization is in the back direction. A *local* monaural sensitivity domain can be identified ±20° around the direction "back". Superior localization acuity for rear locations compared with lateral locations was also reported in [24]. This is not general, but it could be due to the local increase of the monaural sensitivity near to the median plane.

The binaural sensitivity domain can be defined as the overlapping area of the two monaural sensitivity regions. This assumes that the interaural and complex auditory sensitivity is not based only on the monaural sensitivity of the HRTFs. Humans try to face the sound sources for the best localization and use the interaural differences and the binaural fusion. In the median plane no interaural differences appear and only the HTRF should deliver all localization cues. In real-life situations head movements are very useful and important to find the source. If they are not present, front-back confusion and poor localization performance appear.

Similarly, if we can find the most sensitive regions for one ear, we will find the spatial domain where the sensitivity and possibilities of extracting directional information is the worst. This minimum is at ca. 250°-260° in the head-shadow area. Local minimum at –90° in the Interaural Level Differences was also found and modeled by a rigid sphere [25].

3.2 Frequency Limits in the Lateral-Contralateral Evaluation

There are different limited areas in the frequency domain partitioned by "cut-off" frequencies during the elevation of the sound source information.

The limit at 1500-1600 Hz is well known from the literature [2, 9, 26, 27, 28]. The HRTF has five major resonant points: 3, 5, 9, 11 and 13 kHz but there are large individual differences. The high frequency components are responsible for the localization: the sensation is more correlated with the real source direction if the signal has components above 5 kHz. Above 1600 Hz the lateralization is made based on the envelope, below 1600 Hz it is based on Interaural Time Delays (ITD) [29]. Interaural Level Differences (ILD) are present from 20 Hz-20 kHz but they become important above 500 Hz. Monaural spectral features of the pinna appear above 3-3,5 kHz, primary for elevation cues [23]. Low frequency elevation cues are not due to the pinna but to the torso below 3 kHz [30]. We can support this observation, as we did not observe any effects or deviations below 1600 Hz in the HRTFDs.

Shadowing effects cause random incidence. This means, the HRTFs of the contralateral ear vary too rapidly and randomly to evaluate and decode high frequency information and the SNR is less, than on the lateral side. Our test with the torso wearing a baseball cap supports the finding that shadowing and diffraction effects are responsible for the large high frequency deviations in the HRTFs. The frequency, from where these effects will be effective, depends on the azimuth (marked as black areas on Fig. 3-4), on the elevation and on the environment near the head as well. The variations of this "cut-off frequency" are shown on Fig.5 as function of azimuth. This averaged result is calculated from –10° up to +60° elevation for all measurements with the bare and dressed torso. The lowest value of 3 kHz is in the area of the minimum monaural sensitivity supporting the findings in [22, 23, 31].

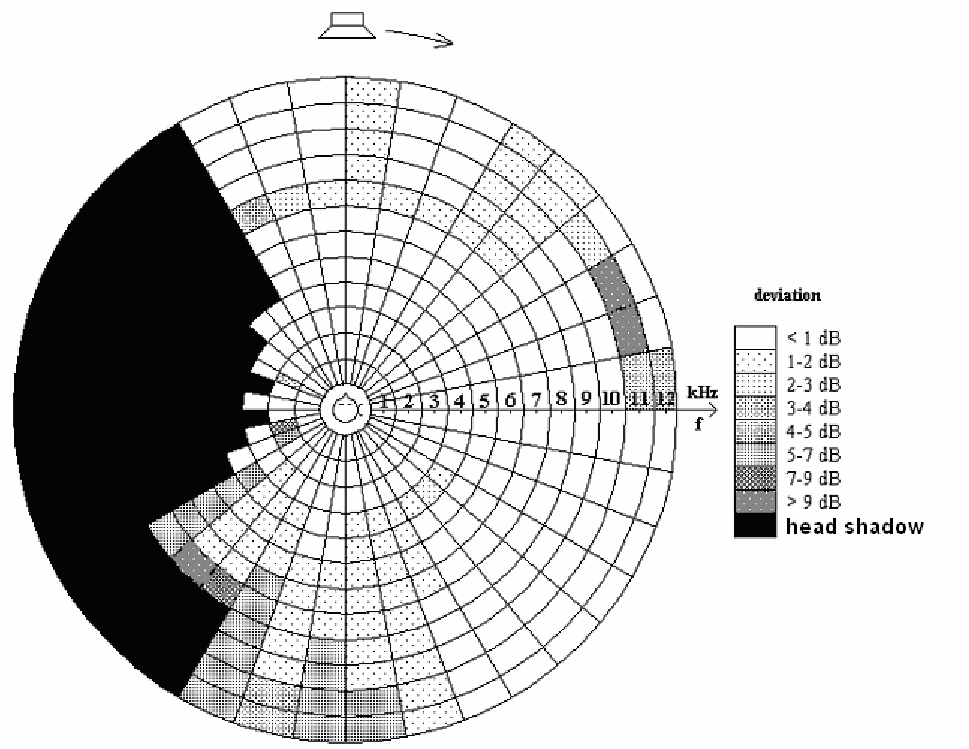

Fig.3. 2D spatial representation of the magnitude of HRTF data for a fixed elevation as function of azimuth and frequency. The polar histogram shows the deviations between HRTFs from repeated measurements in the horizontal plane for the right ear. The natural deviations of the HRTFs caused by the filtering and shadowing effects are shown as unsigned absolute values in dB. The circles correspond to frequency domains with 1 kHz bandwidth marked with the center frequency (linear scale). Note the head-shadow area (filled black) and the domain caused by the pinna between 60 and 90 degrees at 11 kHz and the contralateral side at 2 kHz ("bright spots" [31, 32]). In the head-shadow area HRTFs vary more than 9 dB after independent measurements from the same direction.

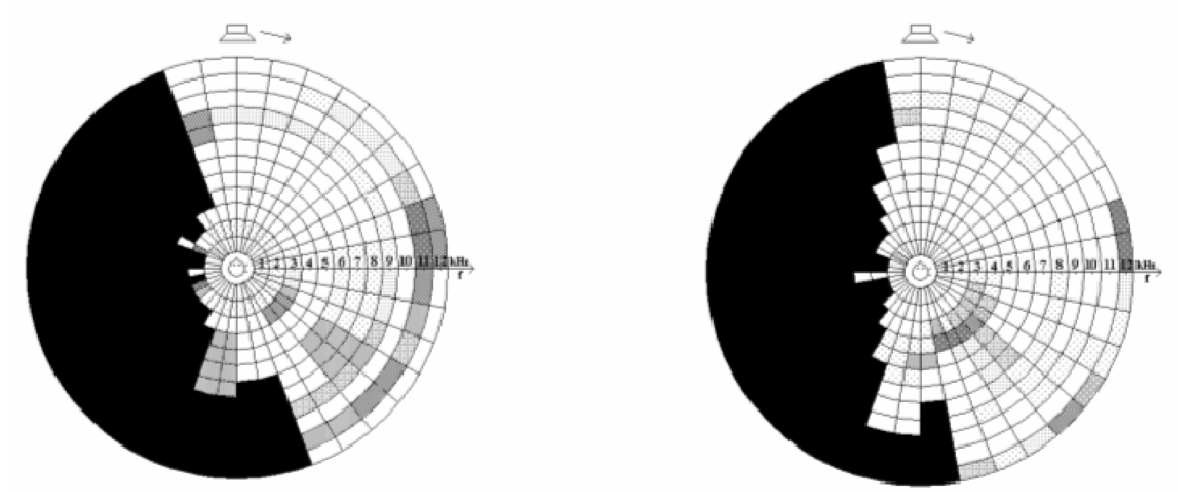

Fig.4. Spatial representation of the HRTF data using a baseball cap on the head of the dummy -head from elevation 10 degrees (left) and 20 degrees (right). Compare with Fig.3. Note the extent of the shadowed area due to the visor of the baseball cap.

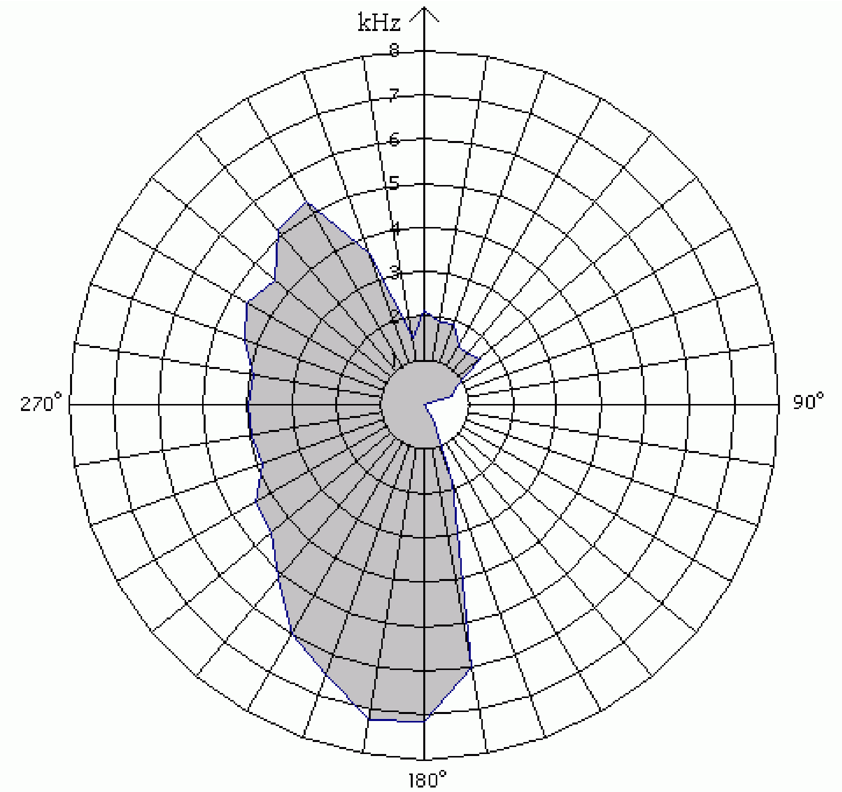

Fig.5. Frequency limit of the head-shadow area as the function of azimuth. Results are averaged over all HRTFDs. The lowest “cut-off frequency” of 3000-3500 Hz is at the minimum of the monaural sensitivity (250-290°).

Although this frequency limit depends on azimuth we can define a stationary value around 3500 Hz. Near to this frequency limit changes and differences in the re-measured HRTFs appear both by the closer and by the contralateral ear. In the shadowed area only some low frequency components will be affected at 1600, 1800, 2200 and 2500 Hz. These so called “bright spots” were found by *Shaw* e.g. at 1,9 and 2,4 kHz [3, 25, 31]. On the other hand, the closer ear will be affected at high frequencies: 9, 11, 4-5 kHz, and only seldom under 3 kHz. Special is the 8 kHz component where the most significant differences appear. At 3500 Hz (speech!) there is evaluation on both the lateral and contralateral side. This is the domain, where neither phase nor intensity differences provide an effective cue (at intermediate frequencies) [26].

This evaluation uses the HRTFDs as a tool for representing the repeatable property. If we re-measure the HRTFs from the same direction without any changes, then we dived them and plot the difference, it has to be a flat (0 dB) line. This is true only in the monaural sensitivity region as shown in Fig.3. These areas are white, and this means, we can re-measure the HRTFs with an accuracy of less than 1 dB. As the sound source is moving on, this repeatable property is getting worse (grey filled blocks) and reaching the head-shadow area (black filled blocks) the re-measured HRTFs even from the same direction have differences more than 10 dB. We assume that this is the natural behaviour and property of the dummy-head (and the real human head as well) due to the shadowing effect of the pinna and head. Fig.4. stated this observation, because the visor of the baseball cap caused the same effect: the extension of the shadowed area. We have some results from the elevation 45 and 60 degrees where less shadowing effect appears and so the extent of the black filled area is decreasing [17-19]. Measured HRTFs from the direction “above” (90 degrees, over the head) are the same and can be measured and re-measured in 1 dB precision. In this case, turning of the head (turntable) is irrelevant for the measurement and all the HRTFs have to be identical. By practical application, this means, we will need less HRTFs and decreased resolution of measured HRTFs as the simulated source is moving up inside the upper hemisphere. We may have a one degree spatial resolution in the horizontal plane (360 HRTFs), but we only need a single HRTF from above.

The effect of shadowing can be regarded as a time-dependent filtering: the filters (HRTFs) from the same direction vary over time, they look different in the high frequency region as we re-measure them later. We made measurements hours and also only seconds later without entering the anechoic room and the HRTFDs showed significant differences. This phenomenon makes the time domain analysis of the incoming signals noisy for the hearing system.

4 BINAURAL EVALUATION

Dummy-heads were evaluated already in listening tests [34-35]. The importance of the HRTFs was obvious during solving the localization tasks. Our investigation does not include (yet) listening test with dummy head’s HRTFs, only spectral analysis was made in the frequency domain. The HRTFDs confirm the important role of the interaural differences. If the source is in the monaural sensitivity region of one ear, the differences and changes due to the environment appear in the high frequency regions. At the same time, the HRTF of the contralateral ear will be influenced at lower frequencies and this result in an increased ILD. We do not find that frequency components vary in the way to decrease the ILD. Diffraction of low frequency components results in amplification on the contralateral side. Head and pinna reflections are responsible for detection and evaluation on the lateral side. The closer ear in the high frequency regions in the monaural sensitivity domain will evaluate the information encoded in the sound waves. The contralateral ear makes evaluation of some low frequency elements where no high frequency information is available (low-pass filtering). Shadowing-effects affect the localization:

it causes random incidence, secondary sound paths, diffuse-like sound field and no primary wave front. Head shadow is the natural reason for that, but objects near the head can influence this phenomenon.

5 SUMMARY

Dummy-heads are well suited for long-term, accurate and standardized HRTF measurements. A precisely controlled system is able to set and re-set the spatial directions, to re-measure and to analyze recorded transfer functions easily. We investigated the role of the HRTFs in human localization through accurate HRTF measurements.

By analyzing the monaural HRTFs of the dummy-head, we can determine the monaural sensitivity regions of the hearing system. We have found this to be symmetrical to the median plane (±20°) and the axe of 45 degrees as the most sensitive monaural direction. The extent of the monaural „antenna footprint" of the ear in the horizontal plane is from about -20 degrees to 90 degrees.

Shadowing effects can be seen as the worst effect: the HRTFs are influenced and they vary too rapidly to allow high-frequency recognition. Re-measured HRTFs in this region vary more than 10 dB above about 2 kHz and they make the incoming signal "noisy". This is a natural phenomenon in the head-shadow area, especially near to the horizontal plane. Above 30 degrees of elevation it is becoming less important. This fact was also supported by the measurement with the baseball cap.

Finally, we tried to determine the spatial extent of this area depending on the environment near the head. The cut-off frequency of the head-shadow area varies from 3-6 kHz, although there seems to be a constant value of about 3500 Hz that separates the lateral-contralateral evaluation.

6 REFERENCES

[1] P. Minnaar, S. K. Olesen, F. Christensen and H. Møller, "Localization with Binaural Recordings from Artificial and Human Heads," J. Audio Eng. Soc., **49**(5), 323-336 (2001).

[2] J. Blauert, *Spatial Hearing* (The MIT Press, MA, 1983).

[3] E. A. G. Shaw, "Transformation of sound pressure level from the free-field to the eardrum in the horizontal plane," J. Acoust. Soc. Am., **56**(6), 1848-1861 (1974).

[4] S. Mehrgart and V. Mellert, "Transformation characteristics of the external human ear," J. Acoust. Soc. Am., **61**(6), 1567-1576 (1977).

[5] D. Hammershøi and H. Møller, "Free-field sound transmission to the external ear; a model and some measurement," *DAGA'91* (1991), pp. 473-476.

[6] C. B. Jensen, M. F. Sorensen, D. Hammershøi and H. Møller, "Head-Related Transfer Functions: Measurements on 40 human subjects," *Proc. of 6th Int. FASE Conference* (1992), pp. 225-228.

[7] H. Møller, M. F. Sorensen, D. Hammershøi and C. B. Jensen, "Head-Related Transfer Functions of human subjects," J. Audio Eng. Soc., **43**(5), 300-321 (1995).

[8] D. Hammershøi and H. Møller, "Sound transmission to and within the human ear canal," J. Acoust. Soc. Am., **100**(1), 408-427 (1996).

[9] W. M. Hartmann, "How we localize sound," Physics Today, **11**, 24-29, (1999).

[10] H. Møller, M. F. Sorensen, C. B. Jensen and D. Hammershøi, "Binaural Technique: Do We Need Individual Recordings?," J. Audio Eng. Soc., **44**(6), 451-469 (1996).

[11] J. C. Middlebrooks, "Narrow-band sound localization related to external ear acoustics," J. Acoust. Soc. Am., **92**, 2607-2624 (1992).

[12] H. Fisher and S. J. Freedman, "The role of the pinna in auditory localization," J. Audiol. Research, **8**, 15-26 (1968).

[13] M. Morimoto and H. Aokata, "Localization cues of sound sources in the upper hemisphere," J.A.S. of Japan, **E 5**, 165-173 (1984).

[14] A. J. Watkins, "Psychoacoustical aspects of synthetized vertical locale cues," J. Acoust. Soc. Am., **63**, 1152-1165 (1978).

[15] R. A. Butler and K. Belendiuk, "Spectral cues utilized in the localization of sound in the median sagittal plane," J. Acoust. Soc. Am., **61**, 1264-1269 (1977).

[16] Gy. Wersényi, "Measurement system upgrading for more precise measuring of the Head-Related Transfer Functions," *Proc. of INTER-NOISE 2000*, Vol. 2, pp. 1173-1176, 2000.

[17] Gy. Wersényi and A. Illényi, "Differences in Dummy-Head HRTFs Caused by the Acoustical Environment Near the Head," Electronic Journal of „Technical Acoustics" (EJTA), http://ejta.org/en/wersenyi1, Russia, **1**, 1-15 (2005).

[18] A. Illényi and Gy. Wersényi, "Evaluation of HRTF data using the Head-Related Transfer Function Differences," *Proceedings of the Forum Acusticum 2005* (2005), pp. 2475-2479.

[19] A. Illényi and Gy. Wersényi, "Environmental Influence on the fine Structure of Dummy-head HRTFs," *Proceedings of the Forum Acusticum 2005* (2005), pp. 2529-2534.

[20] Gy. Wersényi and A. Illényi, "Test Signal Generation and Accuracy of Turntable Control in a Dummy-Head Measurement System," Journal of the Audio Engineering Society, **51**(3), 150-155 (2003).

[21] Gy. Wersényi, "Evaluation of dummy-head HRTFs in the horizontal plane based on the peak-valley structure in one-degree spatial resolution," *Proceedings of INTER-NOISE 2006*, 2006.

[22] C. I. Cheng and G. H. Wakefield, "Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space," J. Audio Eng. Soc., **49**, 231-249 (2001).

[23] V. R. Algazi, C. Avendano and R. O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies," J. Acoust. Soc. Am., **109**, 1100-1122 (2001).

[24] S. E. Boehnke and D. P. Phillips, "Azimuthal tuning of human perceptual channels for sound location," J. Acoust. Soc. Am., **106**(3), 1948-1956 (1999).

[25] D. S. Brungart and W. M. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions," J. Acoust. Soc. Am., **106**(3), 1465-1479 (1999).

[26] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am., **62**, 157-167 (1977).

[27] T. T. Sandel, D. C. Teas, W. E. Feddersen and L. A. Jeffress, "Localization of sound from single and paired sources," J. Acoust. Soc.Am., **27**, 842-852 (1955).

[28] W. Mills, "On the minimum audible angle," J. Acoust. Soc. Am., **30**, 237-246 (1958).

[29] F. L. Wightman and D. J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am., **91**, 1648-1661 (1992).

[30] K. Genuit and H. J. Platte, „Untersuchungen zur Realisation einer richtungsgetreuen Übertragung mit elektroakustischen Mitteln," *DAGA'81* (1981), pp. 629-632.

[31] E. A. G. Shaw, "The external ear," in *Handbook of Sensory Physiology 1*, Auditory System, Anatomy Physiology Ear (Springer, New York, 1974).

[32] F. E. Toole, “In-head localization of acoustic images,” J. Acoust. Soc. Am., **48**, 943-949 (1969).

[33] H. Møller, “On the quality of artificial head recording systems,” *Proceedings of INTER-NOISE 97*, pp. 1139-1142, 1997.

[34] P. Maijala, “Better binaural recordings using the real human head,” *Proceedings of INTER-NOISE 97*, pp. 1135-1138, 1997.

[35] H. Møller, D. Hammershøi, C. B. Jensen and M. F. Sorensen, “Evaluation of artificial heads in listening tests,” J. Acoust. Soc. Am., **47**(3), 83-100 (1999).