Akusztikai Szemle (Akustische Rundschau, Acoustic Review), Vol VI, Nr.2, 2005, pp. 23-28.

On the Amplification-Damping-Effect of the Pinnae and of the Head

A fülkagyló és a fej erősítő-csillapító hatása a külső fül átviteli függvényére

WERSÉNYI György

Széchenyi István University, Department of Telecommunications, H-9026, Győr, Egyetem tér 1., Hungary.

Summary

With the same German title presented Tarnóczy in 1992 his first measurements about the role and effect of the outer ears, head and torso [1]. He used pink noise excitation and a Brüel Kjaer dummy-head in a monaural measurement in the horizontal plane. Ten years later we re-installed and re-measured the HRTFs of the same dummy-head using state-of-the-art measurement equipment with increased spatial resolution and accuracy. It was shown that the environment near the head does have significant influence on the transfer functions and still lots of questions remained.

1 Introduction

Measuring of the Head-Related Transfer Functions has a long history. It began at the 60s and even in the new century we develop new methods and equipment to increase spatial resolution or measurement accuracy [2-7]. It has been proved that the HRTFs – as the first step of evaluation in spatial hearing – play a significant role. As a directional dependent filtering they determine localization cues and thus they help by solving localization problems. Over 50 years of investigations have shown the importance of spatial resolution, individual measurements, or headphone equalization during playback in virtual simulation [8-11]. We got know common problems such as insufficient localization performance, in-the-head localization or front-back confusions. All these lead us to revise this novel investigation of Tarnóczy from 1992 that questioned the role of the fine structure of the HRTFs by focusing on the disturbing influence of the head and torso (including hair and clothing) [12-14].

2 A long time ago in an anechoic chamber far, far away…

In 1992 Tarnóczys' installation included measurement equipment of his age [1]. He placed a BK 4128 head and torso simulator on a turntable in the anechoic room. The turntable control did not allow precise setting of azimuth, so the horizontal plane resolution was 15 degrees. Furthermore, only some preliminary results were obtained in different elevational positions. Monaural measurements were made in the horizontal plane using pink noise excitation. Free field spectrum of the pink noise signal was flat in the 100 Hz – 10000 Hz domain (±5 dB). Relative spectra were calculated as usual in HRTF measurements. Due to the monaural measurement "0 deg" was chosen to be the direction of the axe of the ear. The measurement analyzer was the BK 2133 real time analyzer in 1/12 octave resolution.

3 Return of the dummy-head

Ten years later the same BK dummy-head returned in the anechoic room. We re-installed the measurement system and extended it with recent equipment, upgraded the measurement chain and digital signal processing methods [15, 16]. In order to do this we

- replaced the synchronous motor in the turntable with a stepping motor,
- this motor was controlled through a PC in 1 degree-steps with an accuracy of about 1%,
- elevational settings were made using a laser targeting system from -45 up to +90 degrees with an accuracy of 0,7%,
- specially created pseudo-random white noise was used and 2-channel binaural measurements were made.

Furthermore,

- the DSP board allowed 50 kHz sampling frequency and 16 bit resolution in a 4096 point FFT,
- full automatic control was programmed and big amount of measurement data were collected and evaluated,
- and 89 dB SNR was obtained using 768-times averaging.

Our goal was to completely re-measure the dummy-heads' HRTFs with this precision and to investigate the acoustically relevant space near the head. We did all this in order to increase the measurement accuracy and avoid reflections, permanent errors etc. Others stated and suggested Tarnóczys' results being related to inaccurate measurement setup and measurement errors, first of all due to room reflections. May or may not be, at the reinstallation we reduced uncertain parts and methods, extended and revised the measurement setup and digital signal processing methods that lead us to have a transfer function measurement system with a repeatability of about 0,5 dB in the entire frequency region independent of azimuth and elevation.

Parts and methods which were appropriate	Parts and methods to modify and revise
no measurement in the first period of the input signal (against step-answer appearing)	the 5° step turning of the table is not enough (horizontal plane measurement)
repeated measurements and averaging to avoid random noise effects	the precision and the reproducibility of the stopping positions of the turntable
82 msec. period pseudo-random noise input signal	extension of the elevation under -15°
a precise laser-beam direction calibrator for precise elevation determination	more flexible, and comprehensive measurement software
2 channel, 16 bit resolution, 50 kHz sampling frequency	Reduction of the mains (220 V) disturbing effect

Table 1. Summary of the satisfactory and non-satisfactory elements of the earlier measurement [15].

4 Then and now…

"Die monaurale Untersuchungen ermöglichen nicht nur die spektralverzerrende Wirkung der Ohrmuschel, sondern auch die des Kopfes studieren."

Although we made 2-channel simultaneous measurements, results were first presented for one ear only. Monaural presentation is well suited to show differences between HRTFs from the same direction. According to Tarnóczys' statement, using monaural HRTFs we can study every effect near or on the head that influence the transmission and the fine structure of the HRTFs.

"Es wurde die wichtige Annahme vorausgesetzt (...), dass bei Substraktion zweier Spektren voneinander, diese Komponenten ausfallen werden"

The definition of the HRTFs applies spectral relativity. They are defined as the transfer function from a given direction (in the head-related coordinate system) divided by the reference spectrum measured with an omnidirectional microphone instead of the dummy-head in the same measurement chain:

$$HRTF(j\omega) = \frac{H_{outerears}(j\omega)}{H_{reference}(j\omega)} \tag{1}$$

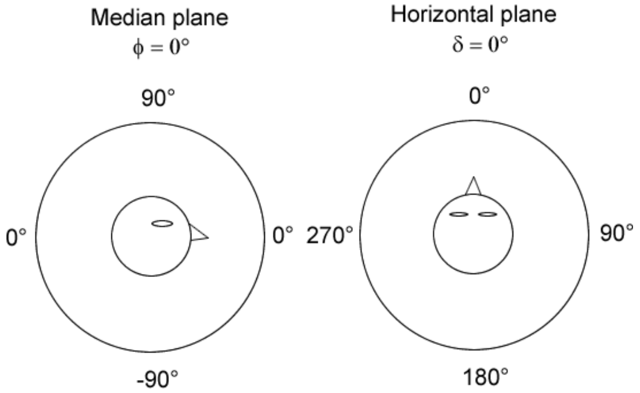

Fig.1. The head-related coordinate system according to our measurement (as usual). Tarnóczys' measurement used 0 degree for the ear-axis in the horizontal plane, thus 180 degrees corresponded to the contralateral ear.

This idea could applied on a HRTF itself: if we devide (or substract in dB) two HRTFs from the same direction we lose individual properties and we are able to determine spectral components that varied. Imagine a dummy-head placed in the anechoic room and a loudspeaker from a given direction radiating white noise. After that, we go in to place clothes, glasses, hair, hat etc. on it and re-measure the HRTFs. Now we divide these spectra, so the "naked" torso is the reference condition, and we plot them [17, 18]. That is what we actually did…

“Wir haben ihn mit einer Jacke und mit einer Perrücke versehen”

In our measurements the objects we have been focused on were: four different kinds of glasses, four different but similar baseball caps and three toupees with different length and haircut. Moreover, some results we obtained from measurements with clothing [19].

“Bei direkten Schalleinfall fand man keine wesentlichen Unterschiede”

We can support Tarnóczys’ observation: the least significant differences appear at the directly radiated ear as the sound source is in the ear-axe.

“Größere unterschiede kamen hervor bei +90 Grad (frontal) 2-4 dB und besonders bei 180 Grad (contralateral ear) 6 dB, wobei die Verschluckung der Kleider und der Haare nach höheren Frequenzen natürlich zunimmt.”

Hair produces a broadband and significant effect, mostly at high frequencies: 9, 10, and 11 kHz. The most important domain is between 4-5 kHz, where the differences are large and permanent as the source is moving in the horizontal plane independent of the elevational position. At lower elevations (up to 20°) the 3,5 kHz components, at higher elevations (above 20°) the 2,5 and the 2,8 kHz components are influenced as well.
At the contralateral ear (in the head-shadow area) differences up to 10 dB appear at 1,8 and 2,2 kHz. Above +30° elevation this effect is less significant.

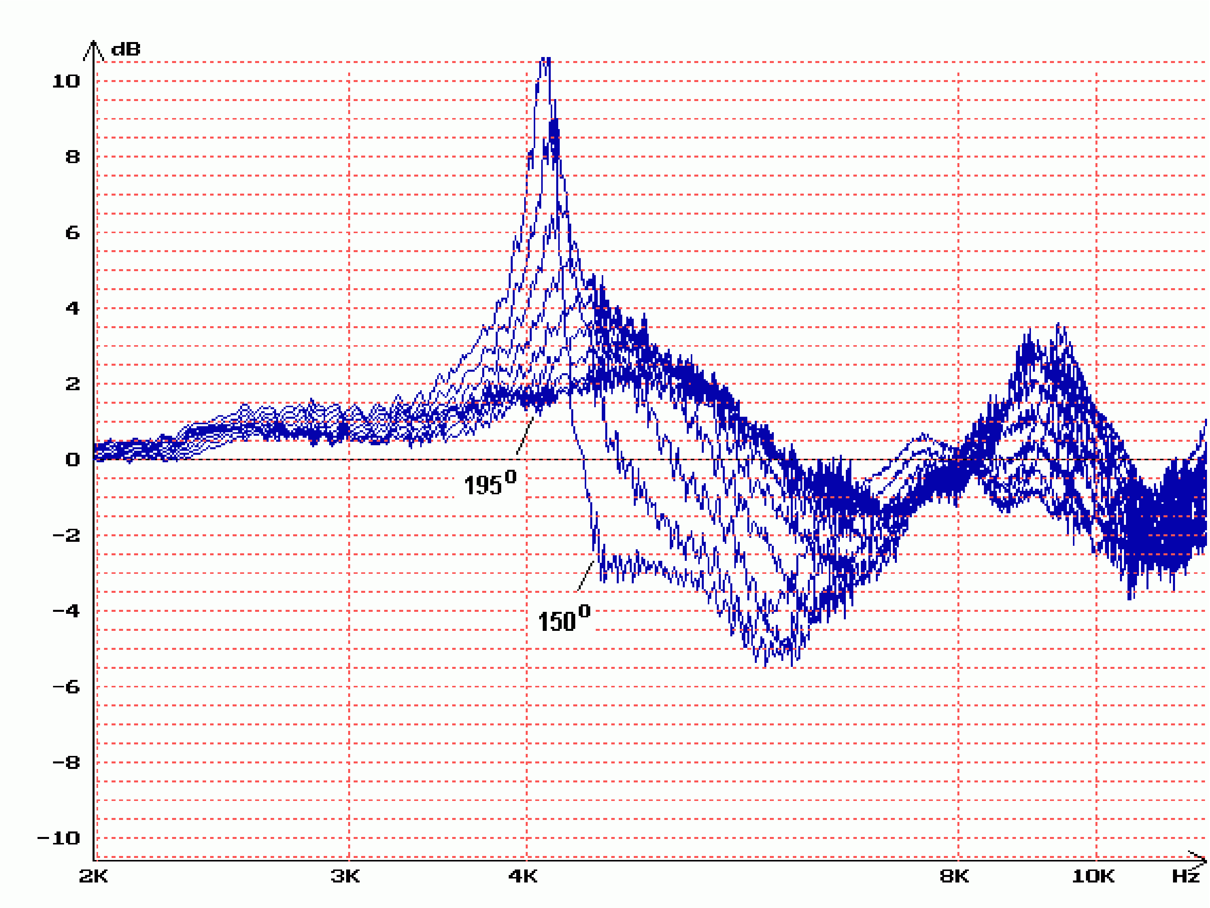

Fig.2. Differences of HRTFs in the horizontal plane as the function of frequency between φ=150°-195° in 5° steps wearing hair.

"Umso mehr eigentümlich ist eine große scharfe Verstärkung bei 180 Grad um 6 kHz, zugunsten des angekleideten Zustandes"

Clothing has a common damping effect due to sound absorption. A thin T-shirt does not influence the transmission, but a thick shirt or coat has a damping up to 2-3 dB at 2-4 kHz, 3 dB at 8 kHz and 2 dB at 11 kHz. In the head shadow area the low frequency components at 1.5, 1.8, and 2.5 kHz show +2 and +4 dB amplification.
Some information can be found about the effect of the torso with and without clothing in [20]. Undressed torso causes sound pressure level increase at the head between 2-5 kHz. In a diffuse-field the fine structure of the torso and the head below 10 kHz is not significant.

"Ich habe darauf gefolgert dass die Effekte bei 3 und 4 kHz unmittelbar von der Ohrmuschel verursacht sind"

Using a dummy-head we are able to remove the artificial pinnae so it is clearly visible how this fact influences the transmission.

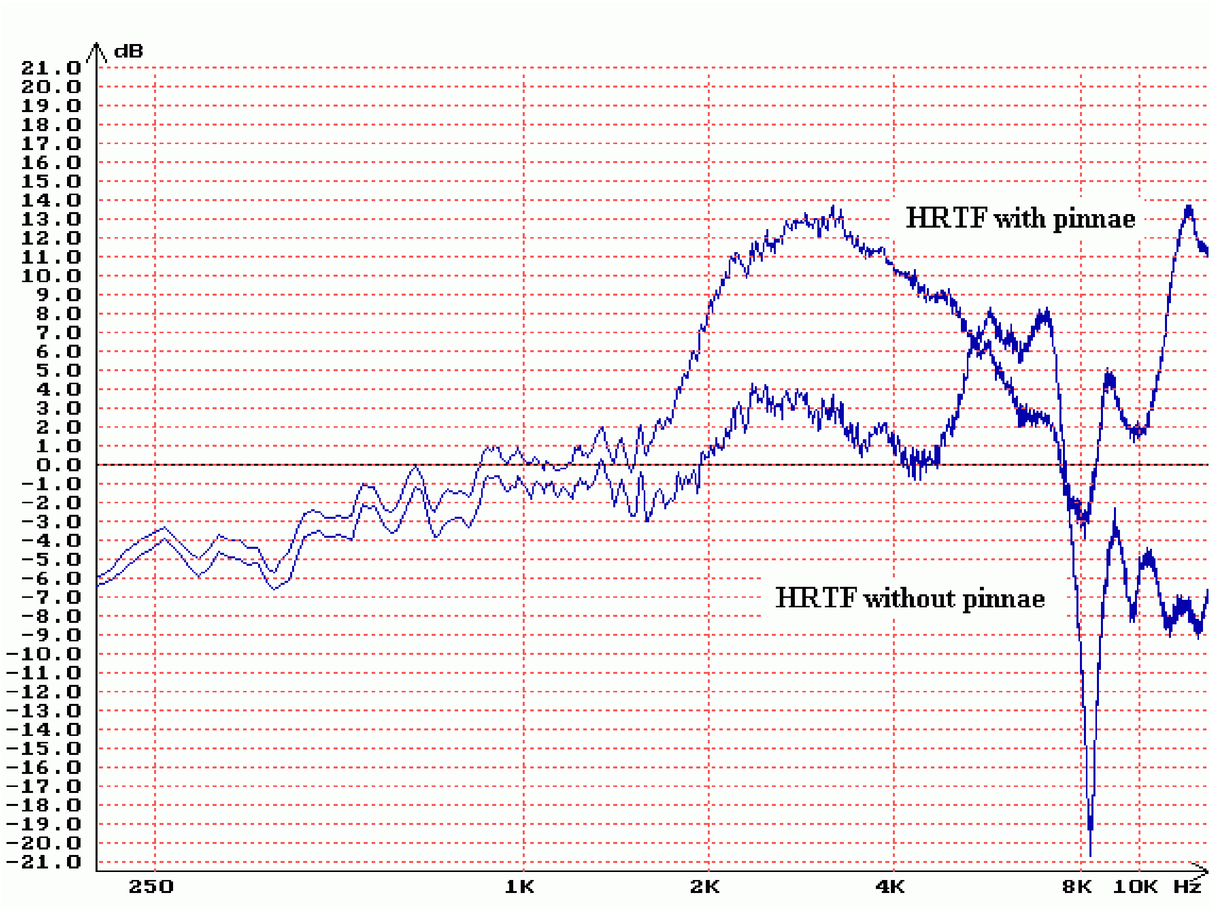

Fig.3. Effect of the pinnae at frontal incidence (φ=δ=0°). Both HRTFs contain the effects of the torso and the head. The reflecting and amplifying effect of the pinnae is clearly visible at the main resonance frequencies of 3, 9 and 11 kHz.

“Man hat auch keinen größeren Effekt erfahren, wenn die Schulter mit einem 5 cm dicken Schaumgummischicht bedeckt war”

We measured large deviation in the φ=60-80° domain around 11 kHz and on the contralateral side at 1600-2500 Hz. The HRTFs of the bare torso are almost identical from this direction in repeated measurements, but between 11 and 12 kHz they are shifted about 25-30 Hz and this difference is enough to affect the difference and produce large deviations [11]. Small shifts in the frequency of sharp notches in the HRTFs from recording to recording result in relatively large variations over very narrow frequency bands [21]. This suggest the importance of the microstructure of the HRTFs. Removing the shoulders or covering them with absorbing materials also decreases this effect but does not eliminate it completely like removing of the pinnae. This high frequency random effect of the pinnae caused by its reflections can be handled by the „multipath-theory” calculating secondary sound paths in the time-domain [22, 23].

"Es wäre also nicht ganz korrekt sehr genaue akustische Daten zu den einzelnen Ohranteilen ankoppeln"

HRTFs are strongly influenced by objects near the listener's head. With accurate measurements we proved that even small changes in the environment cause large deviations in the entire frequency and spatial domain. Thus, the HRTFs can be declared as helpful and basic cue but not as a satisfying element of the localization-decoding procedure. As in every other "information decoding system" they represent a pre-filtering algorithm for higher processing levels but as stand-alone filters they cannot explain the whole decoding procedure.
Objects near the head have different effects on frequency regions on the lateral side and on the contralateral side. Glasses have the smallest effects, because they are thin and cause rather high frequency responses. On the other hand, hair always has influence and caps only in the region where shadowing effects occur (due to the visor). We assume that the most undesired effect for the hearing system is the extending of the shadowed area both in frequency and azimuth, because this can lead to localization errors by losing high frequency information.
Our hearing system seems to be having the ability to "overcome" and disregard some effects appearing in the magnitude responses of the HRTFs without decreasing the localization performance. This feature is deactivated in case of using non-individualized HRTFs and/or headphone playback. The headphone-environment seems to be too "unnatural". This suggests that this "overcome function" of the higher processing system is only active when basic localization requirements are fulfilled. Furthermore, the localization is based not primarily on the magnitude of the HRTFs but on the phase information and higher processing.
All this supports the efforts to increase the artificial recording and binaural playback systems but this is not the same as trying to get better and more accurate HRTFs [24-26].

5 About HRTF reproduction

Basically, there are two different playback situations where the hearing system uses the HRTF pre-filtering. The first is the so called free-field environment. With other words, in the real life situation where no headphone playback is applied. There is no need for HRTF reproduction, we use our individual HRTFs. This means, they are "perfectly accurate" and individual. This is the normal listening situation, we can deliver the best localization performance in listening tests (e.g. the best spatial resolution or the least front-back-confusion rate). The definition of free-field listening requires the lack of reverberation (e.g. an anechoic room) that influences the localization performance.
The other option is to have virtual audio synthesis using headphones. In this "unnatural" environment the free-field listening situation is simulated. In order to do this, we have to measure and record somebody's HRTF set. After that, they have to be reproduced electronically. Furthermore, the transmission chain has to be equalized carefully, first of all the transfer function of the applied headphone. Subjects in this so called virtual listening test deliver worse results in contrast to free-field listening: they report decreased localization performance, in-the-head localization, increased number of front-back reversals, elevation shift etc. These problems are related to the applied HRTFs: individual sets are the best suited for listening test, while dummy-head HRTFs at least. Even randomly selected or an "average" real human heads' HRTFs are better [7-9, 24-26].
It is clearly seen that the HRTFs are one of the important parameters during the localization that influence the measurement results. Our measurement showed that small changes in the environment near the head do really influence the magnitude of the HRTFs: wearing glasses,

having a hair cut, getting dressed or put on a hat affect the HRTF filtering in a wide frequency range depending on azimuth and elevation. In our real life situations (free-field listening) we do not recognize any significant change in our localization performance due to these. This can lead to the conclusion to state the fine structure of the HRTFs not to be that important.

On the other hand, the quality, accuracy, spatial resolution or individuality is much more important during headphone playback. A HRTF set with glasses or with hair can be handled as a particular non-individual HRTF set. Such changes could lead to decreased localization performance.

Out-of-head localization can be achieved using individual HRTFs, reverberation cues, visual cues and head movements (dynamic localization cues). Non-individual HRTFs can be used to generate externally images only when other cues are present [27].

Reverberation, even if it’s only early reflections or attenuated, delayed versions of the direct sound (the non-minimum-phase method), is maybe sufficient to produce external images [28, 29]. Stimuli including reverberation yield lower azimuth errors and higher externalization, but decrease the elevational accuracy.

Except for the interaction of head tracking and HRTFs for azimuth error, there is no clear advantage to using individualized HRTFs for improving localization accuracy, externalisation or reversal rates in virtual synthesis of speech [28]. Results showed no correlation between azimuth error and head size difference. These data may differ when noise stimuli or clicks are used. These effects in the magnitude response let us consider the phase information of the HRTFs to be important. *Time analysis* of the HRTFs has to be performed in the future to see the effect of small head-movements and rapid variations of the HRTFs. HRTFs are dynamic systems, their variations and differences *over time* deliver much more information than the simple magnitude response or the change of the magnitude response caused by changing the source location. Higher processing levels have more influence on the acoustic signal processing during decoding the directional information and are able to distinguish playback situations.

	Free-field Listening	Virtual Audio Synthesis
Individual HRTFs	Always	Seldom (must be measured)
HRTFs from random human, avg. human or dummy-heads	-	Often
Equalization of the headphone etc.	-	Always
Infinite number of applied HRTFs (infinite spatial resolution)	Always	Never
Limited number of simulated HRTFs (finite spatial resolution)	-	Always
Interpolation of missing HRTFs	-	Sometimes
Dynamic changes of HRTFs due to small head-tracking movements	Always (except well fixed head)	Never (must be simulated)
In-the-head localization (error)	No	Yes
Front-back confusion	Seldom	Often
Elevation shift or sources-too-near sensation	Very rare	Often
Existing reverberation	No (in anechoic rooms) Yes (in normal rooms)	No (if not simulated) Seldom (if it is simulated)
Importance of the fine structure and variations of the HRTFs being used	No	Yes
Influence of other localization cues such as head-tracking, reverberation	Yes	Yes

Table 2. Some parameters for the comparison of free-field listening and virtual listening situations.

6 References

[1] T. Tarnóczy: Über den Verstärkerungs-Verminderungs-Effekt der Ohrmuschel und des Kopfes. *Proc. of 6th Int. FASE Conference*, Zürich, pp. 229-232, 1992.
[2] Gy. Békésy: Experiments in hearing. New York, McGraw-Hill book Co., 1960.
[3] Gy. Békésy: Sensory Inhibition. Princetion, NJ, Princton University Press, 1967.
[4] J. Blauert: Spatial Hearing. The MIT Press, MA, 1983.
[5] S. Mehrgart, V. Mellert: Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.* **61(6)**, pp. 1567-1576, 1977.
[6] C. B. Jensen, M. F. Sorensen, D. Hammershøi, H. Møller: Head-Related Transfer Functions: Measurements on 40 human subjects. *Proc. of 6th Int. FASE Conference*, Zürich, pp. 225-228, 1992.
[7] H. Møller, M. F. Sorensen, D. Hammershøi, C. B. Jensen: Head-Related Transfer Functions of human subjects. *J. Audio Eng. Soc.* **43(5)**, pp. 300-321, 1995.
[8] P. Minnaar, S. K. Olesen, F. Christensen, H. Møller: Localization with Binaural Recordings from Artificial and Human Heads. *J. Audio Eng. Soc.* **49(5)**, pp. 323-336, 2001.
[9] H. Møller, M. F. Sorensen, C. B. Jensen, D. Hammershøi: Binaural Technique: Do We Need Individual Recordings? *J. Audio Eng. Soc.* **44(6)**, pp. 451-469, 1996.
[10] H. Møller: Fundamentals of binaural technology. *Applied Acoustics* **36**, pp. 171-218, 1992.
[11] Wersényi, Gy.: *HRTFs in Human Localization: Measurement, Spectral Evaluation and Practical Use in Virtaul Audio Environment*. Ph.D. doctoral thesis, BTU Cottbus, 2002.
[12] P. Berényi, A. Illényi: What does it mean for an HRTF not to have the minimal phase property? *Proceedings of Inter-Noise 96*, Liverpool, pp. 2127-2130, 1996.
[13] Wersényi, Gy., Berényi P.: *Measuring of the Head Related Transfer Functions.* Acoustical Review, Vol. IV., 1.-4., pp. 35-41, 1999, Budapest.
[14] Illényi, A., Wersényi, Gy.: *Discrepancy in binaural tests and in measurements of sound field parameters.* Proceedings of the International Békésy Centenary Conference on hearing and related sciences, 1999, pp. 160-165, Budapest.
[15] Wersényi, Gy.: *Measurement system upgrading for more precise measuring of the Head-Related Transfer Functions*, Proc. of Inter Noise 2000 International Conference, Vol.II., August 27-30, pp. 1173-1176, 2000, Nice, France.
[16] Wersényi, Gy., Illényi, A.: *Test Signal Generation and Accuracy of Turntable Control in a Dummy-Head Measurement System*. Journal of the Audio Engineering Society (AES), Vol.51., No.3, pp. 150-155, March 2003.
[17] Illényi, A., Wersényi, Gy.: *Environmental influence on the fine structure of dummy-head HRTFs.* Proceedings of the Forum Acusticum 2005, 2005, Budapest.
[18] Illényi, A., Wersényi, Gy.: *Evaluation of HRTF data using the Head-Related Transfer Function Differences.* Proceedings of the Forum Acusticum, 2005, Budapest.
[19] Wersényi, Gy., Illényi, A.: *Differences in Dummy-Head HRTFs Caused by the Acoustical Environment Near the Head.* Electronic Journal of „Technical Acoustics" (EJTA), Russia, January 2005 (1).
http://webcenter.ru/~eeaa/ejta/eng/abstracts2005eng/wersenyi1eng.shtml
[20] G. F. Kuhn: The pressure transformation from a diffuse sound field to the external ear and to the body and head surface. *J. Acoust. Soc. Am.* **65 (4)**, pp. 991-1000, 1979.
[21] S. Carlile, D. Pralong: The location-dependent nature of perceptually salient features of the human head-related transfer functions. *J. Acoust. Soc. Am.* **95**, pp. 3445-3459, 1994.
[22] D. J. Haigh: Evidence for generation of multipath localisation cues by human pinnae. *Acustica* **84**, pp. 914-917, 1998.

[23] D. Wright, J. H. Hebrank, B. Wilson: Pinna reflections as cues for localization. *J. Acoust. Soc. Am.* **56**, pp. 957-962, 1974.
[24] H. Møller: On the quality of artificial head recording systems. *Proceedings of Inter-Noise 97*, Budapest, pp. 1139-1142, 1997.
[25] P. Maijala: Better binaural recordings using the real human head,. *Proceedings of Inter-Noise 97*, Budapest, pp. 1135-1138, 1997.
[26] H. Møller, D. Hammershøi, C. B. Jensen, M. F. Sorensen: Evaluation of artificial heads in listening tests. *J. Acoust. Soc. Am.* **47(3)**, pp. 83-100, 1999.
[27] W. G. Gardner: 3-D Audio Using Loudspeakers. Kluwer Academic Publ., Boston, 1998.
[28] D. R. Begault, E. Wenzel, M. Anderson: Direct Comparison of the Impact of Head Tracking Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source. *J. Audio Eng. Soc.* **49(10)**, pp. 904-917, 2001.
[29] J. O. Pickles: An Introduction to the Physiology of Hearing. Academic Press, London, 1982.