Audio Engineering Society

Conference Paper

Presented at the 21st Conference
2002 June 1–3 St. Petersburg, Russia

Sophisticated tube headphones for spatial sound reproduction

Klaus A J Riederer[1] and Risto Niska[2]

[1] *Laboratory of Computational Engineering (LCE) & Laboratory of Acoustics and Audio Signal Processing (ACO) Helsinki University of Technology (HUT), P.O. Box 9400, FIN-02015 HUT, Finland, URL: www.lce.hut.fi/~kar Correspondence should be addressed to the corresponding author (Klaus.Riederer@hut.fi)*

[2] *Unides Design Ay, Sallatunturintie 1 J 75, FIN-00970 Helsinki, Finland*

ABSTRACT

Custom insert-type tube headphones are presented in the study. The sound source is located nearby the human ear canal and therefore the complex problem field of measuring sound sources in the human auditory canal is addressed in detail. The UD-ADU1a headphones have a frequency response of 30-13000 Hz (±10 dB) as measured with a standardized artificial ear, and their upgrade prototype demonstrate a minimal ±5 dB deviation (30-9000 Hz), both 1/3-octave smoothed. The replaceable ear-tip of the headphones attenuates background noise 15-20 dB, allowing a very precise positioning of the sound source. The non-magnetic stimulators are used for reproducing synthetic three-dimensional sounds in neuro- and psychophysiological research.

INTRODUCTION

High-quality headphones allow in principle a good means for realistic three-dimensional (binaural) sound reproduction because of their good channel separation, i.e., due to the lack of cross-talk. Thus the signal processing for the 3-D sound reproduction over headphones is more straightforward and robust than over loudspeakers, as the latter suffer from the "sweet spot" problem, i.e., the limited listening area. Dynamic cues, including unnatural ones, can be incorporated with a head-tracking device connected to the headphones.

For precise perceptual studies of 3-D sound the use of conventional headphones may not be optimal, and it is not even possible for certain cases. The differences in positioning the headphones alter the sound's pathway in the pinnae cavities, creating notable changes in higher frequencies that may affect the perception, see e.g., (Riederer 1998b). Above all, in many brain-imaging methods, such as magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI), only non-magnetic devices can be applied on the subject. However, ordinary headphones applying the dynamic, isodynamic, electret or electromagnetic principle are typically magnetic; electrostatic may not be magnetic, but their large polarization voltage is questionable under the high-energy magnetic field of a fMRI device.

Non-magnetic piezoelectric (insert) headphones are used for metal detectors, but their quality is inferior for fine behavioral experiments, as their frequency response is typically limited to one kilohertz. Earpieces applying other working principles (such as electromagnetic) are typically used in audiometry and telephonometry because of lightweight and good background noise attenuation. Unfortunately, they are magnetic and their sound quality is only mediocre.

Fig. 1. Blocked meatus HRTF measurements, applying a Sennheiser KE 4-211-2 miniature microphone, author as a subject. The arrows indicate the main parts of the pinna; from (Riederer 1998a).

Objective of work

The main goal of this work was to create a personal, non-magnetic, high-quality binaural sound reproduction device for careful psycho- and neurobehavioral experiments. Hence, the principal concern was to obtain maximally flat, wide-range frequency response (approx. 100-10000 Hz), without any active equalization (i.e., applying only passive filtering for the sound source). Because the 3-D auditory stimuli would be synthesized applying HRTFs (both individual and non-individual) with the help of digital signal processing (DSP), the equalization of the reproduction device would be straightforward as long as its raw response would lack any prominently sharp resonance structures. The use of device would require a *personal* type of a sound source, meaning that the sound source would be close to the ears of the listener (headphones / earphones).

Approach of work

Circum-aural and supra-aural headphones can be easily measured with miniature microphones at the entrance of the ear canal (see Fig. 1) while a human subject is listening to them (see, e.g., Møller et. al 1995, Riederer 1998b); the method is not standardized. However, for insert headphones such a simple method cannot be utilized. Consequently, there is no means available to exactly measure the tube (insert) headphones' performance on a human listener. In order to overcome this problem methods approximating the anatomical ear canal structure and acoustical impedance were adopted; on insert earphone measurements see also (Riederer and Backman 1998).

The discrepancy between measurement results and auditory perception was further investigated by applying three different ear canal simulators. The final design of the tube headphones was iteratively developed based on the measures of the earlier designs. The measurements with different ear simulators were performed with different measurement hardware and software. A better resemble to the human auditory resolution (especially notable at higher frequencies) is obtained from the smoothed measurements, not from the raw ones. At the end, the overall quality of the devised tube headphones is compared with different dynamic headphones.

Outline of study

The acoustics of tube headphones is presented in the second chapter. Chapter three presents in detail different artificial ears for insert and other headphone types and the practiced measurements. Results obtained from different types of ear canal replicas are shown with comparisons between various cases. Finally, conclusions are drawn in the fourth chapter.

TUBE HEADPHONES

Based on the previous experience gained on tubular sound transmission lines in the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology (e.g., Airas et al. 1999) and at Unides Design, the solution of using an electromagnetic transducer (loudspeaker) as a sound source and conducting the sound via non-magnetic tubes seemed most profitable for the purpose. Because of the criteria set to the device, a fully new design of the tube headphones was developed.

Construction

The UD-ADU1a tube headphones consist of two discrete, identical units, which can be driven by any normal amplifier yielding at

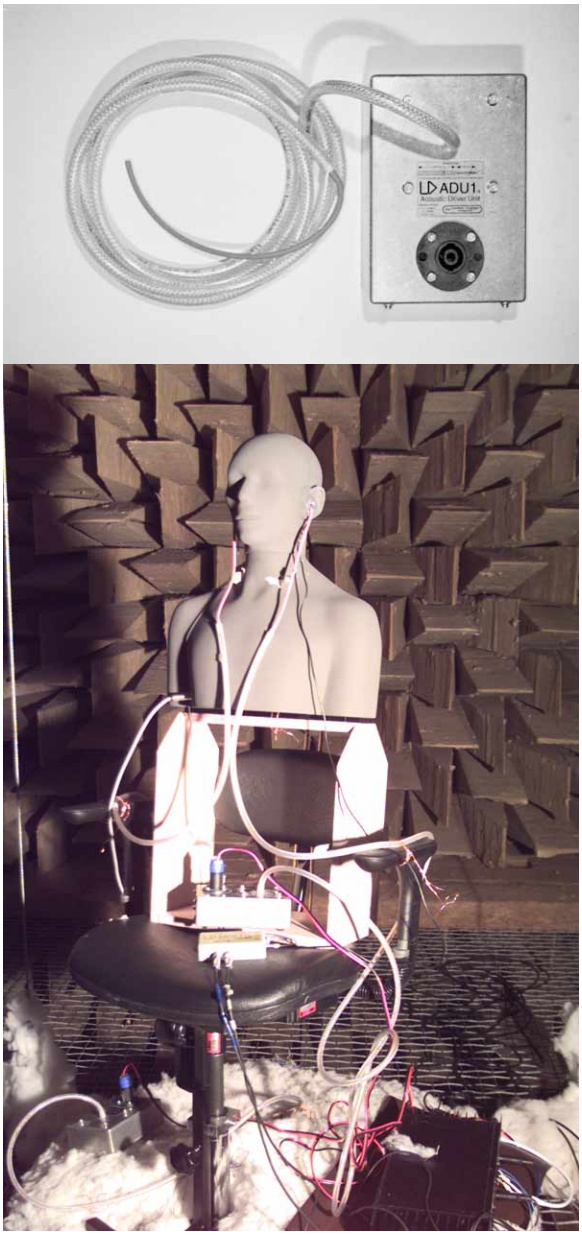

Fig. 2. a) Above: close-up of the UD-ADU1a tube headphones (photo by Risto Niska); b) below: tube headphones attached to the entrance of the ear canals of the used Cortex MK II dummy head (photo by Klaus A J Riederer).

least 100W RMS sound power, see Fig. 2a. The acoustic driver is a modified electrodynamic transducer enclosed in a molded aluminum box, sized 170*120*55 mm^3. To the sound source a 3.00 m long reinforced PVC tube (10 mm in outer diameter, 4 mm inner), to which a smaller laboratory rubber tube of 0.25 m length (5 mm outer diameter, 2.5 mm inner) is fitted. Finally, a replaceable porous EAR-tip (either the larger yellow tip or the smaller light pink one) is attached to the rubber tube. These earplugs typically attenuate the background sound 15-20 dB (depending on the frequency), if properly fitted to the auditory canals. They also allow a precise positioning of the sound source on the subject. Fig. 2b shows the UD-ADU1a tube headphones on the Cortex MK II artificial head (Neutrik 1996) that was used for the measurements discussed in Sections 3.3 and 3.4.

Tube acoustics

The tube headphone acts as an acoustical impedance matching device compensating the discrepancy between the acoustical impedances of the transducer and the ear canal. This is done by the impedance differences between the sound source and the tube areas and ear-tip fitting at the ear canal and also by the passive filtering at the electrical input of the transducer.

One fourth wave transformer

The device may be considered as a tube that is closed from its other end (the tube is tightly fitted to the transducer). This resembles the shape of a human vocal tract, when a neutral vowel is articulated. Standing waves are created in the tube, and at the open end the motion of air is in maximum, but the pressure is in minimum. The antiresonances are located at odd periods of one-fourth wavelengths, i.e., the dips are at frequencies

$$f_n = n * \frac{\nu}{4 * L} \quad , \quad \text{where } n = 1, 3, 5 \ldots \qquad (1)$$

and ν is the speed of sound in air (ca. 343 m/s) and L is the length of the two tubes (see e.g., Rossing 1990). Hence, for the case of $L = L_1 + L_2$ = (3.00 + 0.25 m) = 3.25 m, the dips are located at ca. 30, 80, 130, 180, 240, 290, … Hz. Correspondingly, the resonances (peaks) are located at even periods of one-fourth wavelengths, e.g., at approx. 50, 100, 160, 210, 260, 320, … Hz.

Standing waves can also take place in the auditory canal. This is a common phenomenon in audiometric tests, in which the auricle is cover by tight supra-aural headphones. Now, the ear canal is a tube closed from both ends, and it functions as a ½ wave transformer. The resulting peaks are located half wavelengths apart at both odd and even numbered periods, i.e.:

$$f_n = n * \frac{\nu}{2 * L} \quad , \quad \text{where } n = 1, 2, 3 \ldots \qquad (2)$$

A typical length of 25 mm would yield to a pressure decrease at ca. 6.9 and 13.7 kHz in the audible frequency range (higher modes are above 20 kHz).

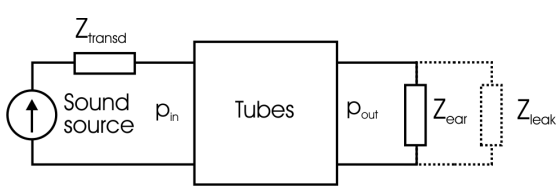

Fig. 3 Simple equivalent circuit of a tube headphone. Here p_{in} is the sound pressure to the PVC tube and p_{out} the sound pressure inside the earpiece, i.e., the EAR-tip.

Ear piece leakage

In order to reproduce low frequencies with a headphone, the acoustic impedance should be low and/or the device should be tightly fitted to the ear(canal) (Poldy 1994). The acoustic impedance of an earphone correlates with the inverse of its size. Circumaural and intra-concha models can be designed to provide a very good seal, whereas intra-aural and supra-aural headphones are rather difficult to make fully sealed; controlled leaks are often used in headphones.

The leak impedance consists of a resistance and an inductance coupled in series. Conversely, the impedance of the ear is capacitive in nature, and hence the leak provides smaller impedance at lower frequencies than the ear (Poldy 1994). In case there is a leakage between the earpiece (EAR-tip) of the tube headphone and the human ear canal, the impedance of leakage couples in parallel to the ear canal impedance; see Fig. 3. This results in a level decrease at lower frequencies. See studies by Rappaport and Spraque (1952) and Riederer and Backman (1998) for further details on a similar problem, i.e., on the improper fitting of stethoscope earpieces.

Equivalent circuit of tube headphone

The acoustical function of a tube headphone can be depicted by an equivalent circuit, see Fig. 3. The tubes can be depicted as a four-pole *transmission line*, terminated by the EAR-tip / ear impedance (Z_{ear}) and a possible leak (Z_{leak}). The transfer function from p_{in} to p_{out} is determined by the acoustical load and the properties of the driver (transducer impedance, Z_{trans}). Moreover, the propagation losses in the transmission line are important; the high frequencies attenuate faster than the low ones. Now, the puzzle is how to fit the driver to the transmission line, so that no reflections from the other (ear) end would occur.

Tube reflections

In case the tube (i.e., transmission line) is not perfectly fitted to ear canal, reflections are seen at the open end (ear canal). They would appear in multiples of the double distance of the tube's length, i.e. their time delay τ would be

$$\tau = n\frac{2L}{\nu} = n\frac{2(L_1 + L_2)}{\nu} \quad , \quad \text{where } n = 1, 2, 3 \ldots \qquad (3)$$

According to Eq. (3), the reflections of a 3.25 m long tube would be delayed 19, 38, 57, ... ms in time.

EAR SIMULATOR MEASUREMENTS

The ITU-T recommendation P.57 (ITU-T 1993) defines three types of *artificial ears* that are devices consisting of an acoustic coupler and a calibration microphone for the measurement of sound pressure in order to calibrate earphones. Type 1 (IEC 318) should be used for *telephone bandwidth* measurements on supra-aural and supra-concha earphones, is it is *not suitable* for low acoustic impedance earphones. *Type 2* (IEC 711) is to be applied on sealed and unsealed *insert earphone* measurements, and its output sound pressure is referred to the *eardrum reference point* (DRP), not at the *ear reference point* (ERP). Numerical 1/3 and 1/12 octave corrections are given to convert the measurements from DRP to ERP, which is the reference sound source point for supra-aural headphones. Type 3 Artificial Ear consists of the occluded-ear simulator IEC 711 added with the ear canal extension terminated with a pinna simulation device. The three recommended pinna simulators, i.e., Type 3.1 Concha bottom simulator, Type 3.2 Simplified pinna simulator and Type 3.3 Anatomically shaped pinna simulator, provide suitable arrangements for measuring different transducer types.

The standardized artificial ears are rough average estimates; they have overall acoustic impedance similar to an average human ear over a limited frequency range. It is obvious that the measurements only approximate the sound pressure at the eardrum, as no real-head probe recording would be feasible. To investigate further the divergence of ear simulators, three types of models were applied: Brüel&Kjær Artificial Ear with the Ear Canal simulator, custom ear canal replicas and Cortex MK II dummy head with anatomically shaped pinna. Smoothed responses give a closer resemblance of the hearing percept, i.e., the human auditory resolution. Iterative measurements were used to improve the design of the tube headphones.

In order to have a high frequency resolution, long FFTs (16384 bins) were used for calculating the frequency responses applying a rectangular window. Traditional (one-third) octave smoothing was created by averaging the results with a moving (rectangular) window. The measurements were unequalized (disregarding the NCF1 Cortex microphone preamplifier / filter, see below), i.e., no equalization for the microphones or the artificial ears was done.

The insert-phone measurements apply an ear canal replica, and thus the SPL obtained is referred *at the eardrum*. This means that calculated frequency responses above 1 kHz differ greatly from other reference point measurements, and they should not be compared directly with, for example, free-field responses of loudspeakers.

Artificial Ear measurements

The tube headphones were measured with a Brüel&Kjær Type 4157 Artificial Ear including a microphone Preamplifier Type 2639 that is individually calibrated for the 1/2“ microphone Type 4134. The cavity in the replica is 7.6 mm in diameter and 8.9 mm length, following the IEC 711 Type 2 artificial ear (cavity dimensions Ø = 7.5±0.1 mm, l = 8.8 ± 0.1 mm). The capsule is screwed (*without* the Protection Grid 3420) against a thin plastic sealing to a round-turned cone-construction that constructs the larger microphone diaphragm (Ø = 13 mm) to the smaller ear cavity (Ø = 7.6 mm). Inside the cavity (2.9 mm from the outer edge) there is a protective metal grill with a mesh size of ca. 1 mm^2.

To the Artificial Ear a Brüel&Kjær Type DB2012 ear canal adapter (shape of a constricted cone, l_1 = 5.0 mm, $Ø_1$ = 7.6 mm; l_2 = 8.6 mm, $Ø_2$ = 13.2 mm) was attached, thus constituting the ear canal simulator a constricted cone with a total length of 22.5 mm (l_1 = 13.9 mm, $Ø_1$ = 7.6 mm; l_2 = 8.6 mm, $Ø_2$ = 13.2 mm). These dimensions are in alignment of a median human ear canal size but the material (stainless steel) does not come close to the human tissue; see, e.g., Blauert 1997.

Equipment

The tube headphones were measured with the commercial MLSSA 10.0 software/hardware (Rife 1996) running in a 266 MHz Windows98 PC. The 32767 long MLS (maximum length sequence) excitation was reinforced with a Yamaha Natural Sound Amplifier MX70 to a sufficient sound pressure level, i.e., ca. 90 dB (A), in order to obtain a good signal-to-noise ratio (SNR). The headphone’s yellow EAR-tip was inserted to the Artificial Ear Brüel&Kjær Type 4157 fitted with the ear canal adapter Brüel&Kjær Type DB2012 so that an acoustical ear canal impedance approximating the average human’s corresponding was applied.

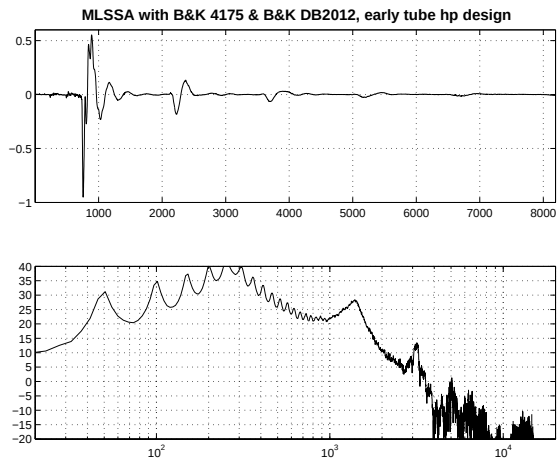

Fig 4. Early developmental version of the tube headphones measured with a Type 2 Artificial Ear and MLSSA-software/hardware; a) Above: impulse response, b) frequency response (unsmoothed).

Methods

Special care was taken in order to avoid leaks between the used yellow EAR-tip (uncompressed l = 13 mm, Ø = 13 mm) and ear canal adapter, and their fastening was secured by moldable silicon putty and adhesive tape. The earplug was put flush to the outer edge of the adapter, i.e., filling ca. 12 mm of the ear canal. The results were obtained as impulse responses by the MLSSA software, and the files were transferred to Matlab5.3/Matlab6 for detailed analysis and visualization. Different designs of the tube headphones were measured, and also the effects such as twisting the tubes and averaging the measurement responses were investigated.

Results

Figs. 4ab demonstrate the impulse and frequency response of an early developmental version of the tube headphone without any filtering for the transducer. The impulse response in Fig. 4a shows the reflections, i.e., the regularly existing antiresonances at ca. 755, 2225, 3700, 5170 samples, τ_s. Thus, the time delay between them is approx.

$$\tau = \frac{\tau_s}{f_s} = \frac{1470}{75472}\,\mathrm{s} = 19.5 \quad \mathrm{ms} \quad \text{, where} \tag{4}$$

f_s denotes for the sampling frequency (f_s = 75472 Hz). This matches perfectly with the calculus in Section 2.2.4 (Eq. 3), and hence these dips are one-fourth wave reflections. Because these standing waves occur at low frequencies, they span over a long duration in the impulse response. A closer look proves the above: cropping the impulse at 2000 samples cancels out the resonancies in the frequency response.

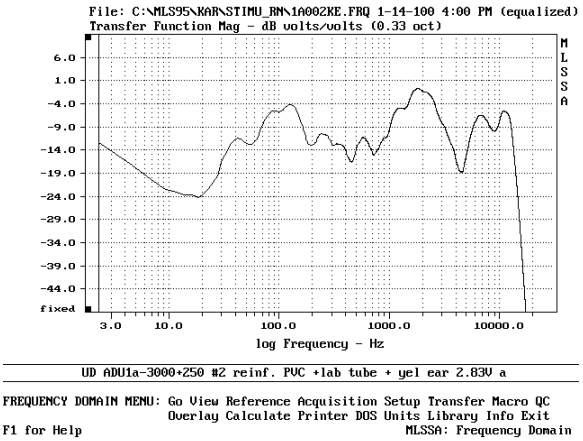

Fig 5. Frequency response of the UD-ADU1a tube headphones measured with a Type 2 Artificial Ear and MLSSA-software/hardware (screenshot from the software), smoothed 0.33 octave.

On the other hand, at higher frequencies the propagating reflections superimpose and interfere with each other, hence the discovered frequency response has somewhat intense resonance structure. The peaks at ca. 250, 1300 and 3100 Hz in the frequency response (rectangular window, FFTsize 16384 bins) in Fig. 4b are created by the peaks in the impulse response at ca. 1000-1500 samples. The transducer is not fitted to the tubes that act as a bad transmission line: high frequencies are strongly attenuated, e.g., above 10 kHz the SPL is –70 dB compared to the maximum level at ca. 200 Hz. Such a strong correction of dynamics would be infeasible by purely DSP means – it would require 12 bits, leaving no sufficient headroom for the actual signal dynamics.

The frequency response of the final version of the tube headphones (UD-ADU1a) shows that the standing waves do not exist any longer, see Fig. 5. The frequency range is lifted to a wide range of 30-13000 Hz (±10 dB). The slightly smoothed (0.03 octave) plot shows three sharper dips, but they yield no problems for the straightforward DSP equalization applied in the 3-D sound reproduction. However, for equalization it is much better to apply more pronounced smoothing. In that way the equalized device is less susceptible to the apparent deviation between the used equalization response and the transfer function caused by the *individual ear canal shape* and typically somewhat inconsistent EAR-tip positioning. The human auditory resolution also motivates the use of stronger smoothing, such as traditionally used 1/3 octave smoothing (or auditory based frequency warping; see, e.g., Smith and Abel 1995).

The measurements also proved that twists in the PVC tube as small as 10 cm in diameter did not affect the frequency response of UD-ADU1a, neither averaging the responses (MLSSA's online averaging; 'go once' vs. 'go 20 times') had any significant meaning.

Custom ear replica measurements

The fixed length and yet debatable material (stainless steel) of the Brüel&Kjær Artificial Ear led the authors to try another more natural-like ear canal replicas using nylon tube, which could also yield a result closer to the listening percept. The effect of the 'ear canal' length was studied by executing different lengths for the replica.

Equipment and methods

The measurements were carried out by a custom-made FFT-based measurement software (Unides Design A-Lab) running in a 300 MHz PC under Windows98 and applying a modified and equalized Soundblaster 128 soundcard and a custom-made battery operated microphone preamplifier (UDMP10D, Unides Design 1999). The stimulus applied was a 16384 (2^{14}) sample long single impulse.

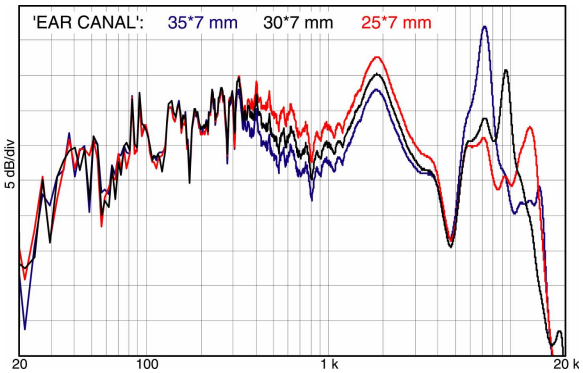

Fig. 6. Frequency responses of the tube headphones with different length (from top to bottom at 1 kHz: 25, 30 and 35 mm) ear canal replicas of a nylon tube (Ø = 7 mm), smoothed 0.1 octave.

The ear canal replica was manufactured from a stiff nylon tube (inner diameter 7 mm, outer 10 mm). A miniature electret microphone Sennheiser KE4-211-2 (Sennheiser 1997) with a 4.75 mm diameter was tightly sealed with an elastic silicone tube to the nylon tube. First, different 'ear canal' lengths were investigated: 25, 30 or 35 mm (constant 7 mm inner diameter), after which the 30 mm length chosen for the reference length. In all cases a light pink EAR-tip (uncompressed l = 13 mm, Ø = 9 mm) ending the tube headphone was carefully placed at the open end of the replica, filling ca. 12 mm of its length.

Results

Fig. 6 demonstrates the early comparative measurements with different 'ear canal' lengths. The petite jitter at low frequencies below ca. 1 kHz is caused by the measurement surroundings and method: the measures were done in a normal room susceptible to traffic and other background noise while applying a single impulse. The curves are slightly smoothed (1/10 octave), in order to better demonstrate the changes between the cases. Below 300 Hz there is no notable variation. At 300 - 4000 Hz the curves demonstrate a very similar resonance structure, the only difference is that the longer distance there is to the microphone, the lower level is obtained. [Note: the sound's near-field in a small closed cavity needs not to follow any power law, Uosukainen (2002).] Above 4 kHz the standing waves in the 'ear canal' interfere with the tube headphones, which causes sharp resonances shifted in frequency. In this case the earplug shortens the effective length of the ear canal by ca. 12 mm, and this effect must be accounted. Using Eq. (2), one get the following approximate values for the first resonance modes: 7.5 kHz (L= 35-12 = 23 mm), 9.5 kHz (L= 30-12 = 18 mm) and 13.2 kHz (L= 25-12 = 13 mm), which are excellently aligned with the measurement results.

The curves demonstrate a crude approximation of the frequency response of the tube headphones inside an average human ear canal. For the authors, the 30 mm response came closest to the auditory percept, hence this length was used for further measurements.

Dummy head measurements

Dummy heads (*artificial heads*, *head and torso simulators*) mimic the anatomy of a median human head (and torso), and apply an elastic *pinna*. Many heads follow the IEC 959 standard for the *pinna shape*. However, the implementation of the *auditory canal simulation* varies in different models, and it might not follow the

ITU-T recommendation of Type 3.3 Artificial Ear. If the dummy head is used for spatial sound studies this is acceptable, because the ear canal itself or the actual placement of the microphones inside the ear canal does not matter when investigating the direction-dependent auditory information (see, e.g., Wightman and Kistler 1989, Hammershøi and Møller 1996). Nevertheless, the ear canal dimensions certainly affect pressure-field (such as insert earphone) measurements; see also (Blauert 1997, p. 302).

The ITU-T recommendation P.57 (ITU-T 1993) Type 3.1 Artificial Ear simulator has a flat plate terminator in the ear canal extension (creating a lengthening of 2.2 mm to the ear canal), and it is recommended for intra-concha earphones sitting on the bottom of the concha cavity. Type 3.2 simulates the pinna by a cavity (Ø = 7.5±0.1 mm, l = 8.8±0.1 mm) terminating the ear canal extension and adding 2.2 mm to the length. It is intended for sealed and unsealed supra-aural and supra-concha earphones up to 8 kHz and measurements on low acoustic impedance earphones. Type 3.3 Artificial Ear (IEC 959) is terminated by a high-quality standard hardness elastomer (ISO 868), whose shape mimics the median human pinna shape. It is recommended to be used for supra-concha earphones not fitting to the Type 1 or Type 3.2 Artificial ears and for intra-concha earphones not intended for sitting on the bottom of the concha cavity.

Disregarding the recommendation, the UD-ADU1a tube headphones (insert type headphones) were measured using a dummy head with an anatomical pinna, following the IEC 959 standard shape. The reason was to study the differences of frequency responses between various types of artificial ears and their relation to the auditory percept. Another reason was to apply the same measurement approach suitable for supra-aural and circum-aural headphones (see also Riederer 1998b), so that the frequency responses between different headphones types could be approximately compared.

Equipment

The UDADU1a tube headphones were measured with custom-made software written in Macintosh Common Lisp, using an object-oriented DSP environment, QuickSig (Karjalainen 1990). The tube headphones were attached to the ear canals of the pinnae (IEC 959) of a Cortex MK II dummy head (IEC 1990), applying the built-in microphones Brüel&Kjær Type 4190 and their reference preamplifier/digital filter unit NCF1 (Neutrik 2002), see Fig. 2b.

The detachable pinna of the artificial head is molded into a hollow aluminum cylinder (ca. l = 30 mm, Ø = 50 mm) that is filled with elastic silicone containing the silicone / brass pinna & ear canal replica, which is then attached to the 1/2" microphone free-field condenser microphone (*with* the Protection Grid 3420). The outer ear's visible ear canal (ca. l = 11 mm, Ø = 10 mm) ends to a dense grill, behind of which is a brass tube (ca. l_1 = 9 mm, Ø = 12 mm, l_2 = 7 mm, Ø = 13.2 mm). The pinna / ear canal system is pressed into the brass tube that lacks an elastic sealing, and this makes the whole construction sensitive to leaks. The ear canal is in total 20 mm long widening to the 'eardrum' end (10 → 12 mm), which establishes a rather poor replica of the real human auditory canal. Subsequently, the acoustical transfer functions of the dummy head's head and ear canals are further modified (equalized) by DSP means in the NCF1 filter unit; altering the system's high frequency response.

The measurement computer (Apple Macintosh Quadra 950) created the stimulus that was reinforced by a Quad 240 power amplifier to ca. 90 dB (A-weighted) SPL. An NB-A2100 acquisition board recorded the NCF1's (linear) analogue output with a 48 kHz sampling frequency and the QuickSig software saved the results directly in Matlab's *.mat* -format. The NB-A2100's high-quality 16-bit Sigma-delta A/D D/A converters were used to buffer the data flow to/from the measurement computer.

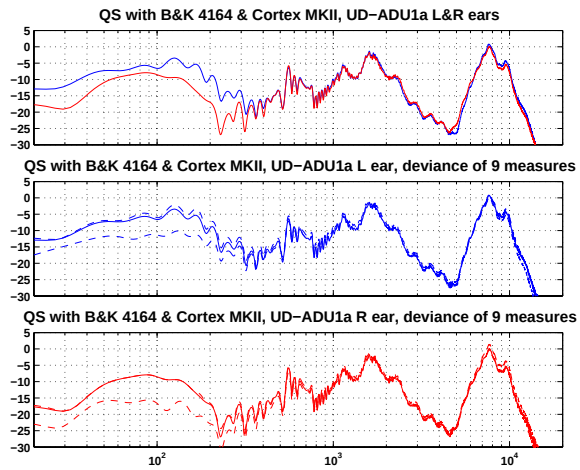

Fig 7. Tube headphone measurements on the Cortex MK II dummy head applying built-in microphones, unsmoothed frequency responses; a) Top: left and right ears; b) middle: left ear, total variation between nine repetitions, c) below: right ear, total variation between nine repetitions.

The stimulus applied was a *random-phase flat spectrum* (*RPFS*) sequence having a flat spectrum in certain time window and frequency points, which can be easily produced using FFT/IFFT techniques, if the sequence length is an exponent of two. Now, a length of 2^{11} (2048 samples) was used. The stimulus was further modified to consist of two subparts with opposite polarities, so that in the IFFT the summing of these parts removed the unwanted DC offset errors (but not the DC component) from the obtained system response. The measurement computer setup was the same as in the HRTF measurement system devised by the author (Riederer 1998ab).

Method

Extensive care was taken so that leaks between the earplug and ear canal would be avoided. The yellow EAR-tip was carefully attached to the ear canal of the dummy head. The rubber tube was held by hand as long it took the squeezed earplug to get fully expanded inside the ear canal, after with the PVC tube was fixed with adhesive tape to the neck of the dummy head. Following the same procedure, ten successive repetition measurements were performed, changing the earplug to a new one after the first five measures. Before each measurement, the EAR-tip was treated so that it was fully expanded before the next insertion procedure. The two ears were prepared one after the other, but they were measured simultaneously.

Results

Fig. 7 shows the Cortex MK II measurements of the tube headphones. Fig. 7a (top) demonstrate petite resonances (5 dB peaks at 50 Hz intervals) at 100-1000 Hz that are caused by standing waves in the headphone tubes. This is because the dummy head's ear canal acoustic impedance differs from the replica's corresponding that was used for designing the tube headphones. Above ca. 6 kHz the preamplifier unit NCF1 raises the amplitude of the measured signal; the frequency response of the headphones is boosted accordingly.

Figs. 7bc prove the excellent repeatability (above 400 Hz) in positioning the tube headphones. The variation between the frequency responses of nine successive measures is less than ±1.5 dB, typically less than ±0.5 dB. [One out ten was corrupted due to a human error — see (Riederer and Karjalainen 1998) for a report on the complexity of transfer function measurements]. The low-

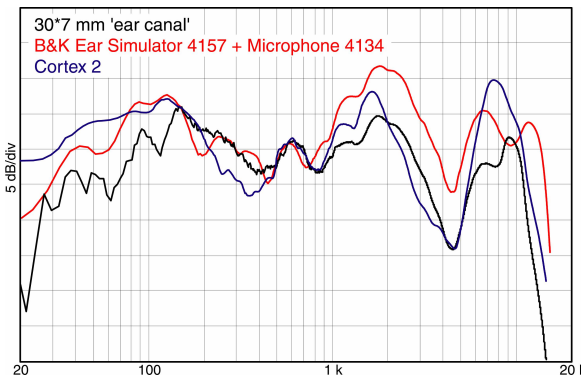

Fig. 8. Comparison on the tube headphone measurements applying different ear canal replicas (at 1 kHz, from highest to lowest level: B&K 4157, Cortex MK II and 30 mm custom replica.

frequency repeatability is still very acceptable, ±3.5 dB for both ears. The reason for the higher variability is obvious: despite the careful insertion procedure, the occlusion of the ear canal has not been perfect, and leaks have occurred (see Section 2.2.2). Because of the protective grill, the EAR-tips fill only ¾ of their length in the dummy head's visible ear canal and have a tendency of not keeping there properly. The human ear canal is longer and narrower than the dummy head's corresponding. Thus the earplugs can mounted more properly, and leaks with human listeners are well avoided.

Comparison of different measurements

There are a number of matters that differ in the tube headphone measurements explained before, which complicates their direct comparison. The Cortex MK II built-in microphones Brüel&Kjær Type 4190 (Ø = 1/2") are free-field condenser microphones (*with* the Protection Grid 3420) and thus their response attenuates above 2 kHz ca. 1-2 dB per octave in pressure-fields. On the other hand, both Brüel&Kjær Type 4134 (Ø = 1/2") and Sennheiser KE4-211-2 (Ø = 4.75 mm) models are condenser pressure microphones and despite the latter's small size it is well insensitive to structure-born vibration due to its back electret design. The mounting of the microphone capsule to the air cavity in the ear canal is executed in various ways, which affects both high and low frequencies.

The material (hardness) of the ear canal replicas varies: stainless steel, nylon tube, and brass-elastomer. Finally, the shapes of the ear canals are different. The cross-sectional area of the auditory canal becomes important at high frequencies, the today's common opinion of that threshold is 5 kHz (Blauert 1997, p. 302).

Now, understanding the differences between the approaches, a direct comparison between the previous ear canal simulator measurements is presented. Ultimately, the tube headphones are evaluated against different circum-aural headphones.

Different ear replica measurements

Fig. 8 compares the previously explained methods, smoothed 1/3 octave. The different ear canal / pinna simulators yield to quite different results, even though the frequency responses demonstrate roughly the same resonance structure. Below 1 kHz the differences are still minor (±5 dB variation), especially the custom 'ear canal' illustrates lower levels. The custom replica being closest to the human physiology is softer than the other replicas (nylon vs. metal) and has no extra air cavity in front of the microphone diaphragm that might introduce the pressure increase.

Above 1 kHz the Cortex and the 30 mm nylon replica measures exhibit fairly similar results — considering that the 5-10 dB boost

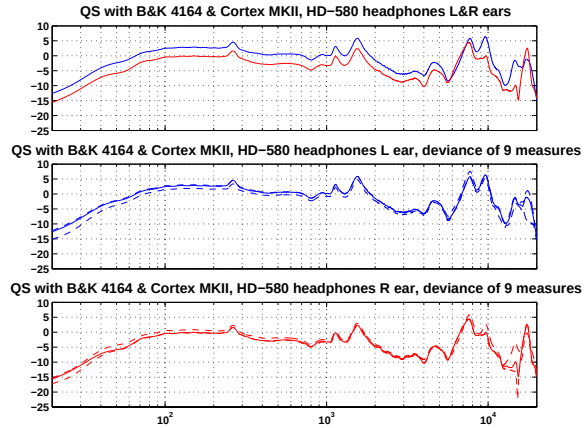

Fig. 9. High-quality circum-aural headphones (Sennheiser HD 580) measured on the Cortex MK II applying the built-in microphones, no smoothing; a) Top: left and right ears, b) middle: left ear, total variation between nine repetitions, c) below: right ear, total variation between nine repetitions.

at the 6-10 kHz range in the Cortex responses is caused by the DSP equalization in the NCF1 preamplifier. The minor diversity between 1-4 kHz between the two responses is well explained by the ear canals' length difference (20 mm vs. 30 mm) as was explained in Section 3.2.2. Between 1-7 kHz the Brüel&Kjær Artificial Ear response is exactly like the custom replicas corresponding, but raised by 8 dB. Also here, the variation in length (22.5 mm vs. 30 mm) causes a level difference. Above 7 kHz the Brüel&Kjær yields to a very different curve, which is most likely caused by the disssimilar geometry: the cone shape of the Brüel&Kjær ear canal (and possibly the microphone mounting).

Circum-aural and supra-aural headphones

The author has investigated in detail the repeatability of headphone-to-(real-)ear transfer functions (PTFs) with circum-aural and supra-aural headphones (Riederer 1998b) that also reveals the typical frequency response of such headphones measured at the *ear reference point*. However, for insert headphones such an approach is not applicable. For this purpose, various headphones were measured following the method presented in Section 3.3.

Fig. 9a shows the unsmoothed frequency response of high-quality circum-aural headphones (Sennheiser HD 580) measured with the Cortex MK II artificial head with the anatomically shaped pinna and built-in Brüel&Kjær Type 4190 microphones coupled to the NCF1 preamplifier / filter. The overall frequency response is very flat at the range of 70-2000 Hz, disregarding the three soft ear canal resonances (compare with the flat responses in (Riederer 1998b), Figs. 38 and 39).

Above 2 kHz the response fluctuates more, ca. ±8 dB. Fig. 9bc reveals the variance between nine successive measures, smoothed 1/3 octave. The repeatability is excellent (±1 dB) up to approx. 12 kHz, after which the petite misplacements of the headphones raise it up to ±4 dB (without smooting up to ±10 dB). Taken together, applying immovable built-in microphones with a dummy head, such kind of results from high-quality devices are expected. However, with real human heads the curves yield to a much larger variation (see, e.g., Møller et. al 1995, and slightly better results in Riederer 1998b).

The inexpensive model (Sennheiser HD 570) is presented in Fig. 10 with very appreciated supra-aural headphones, Grado SR-80,

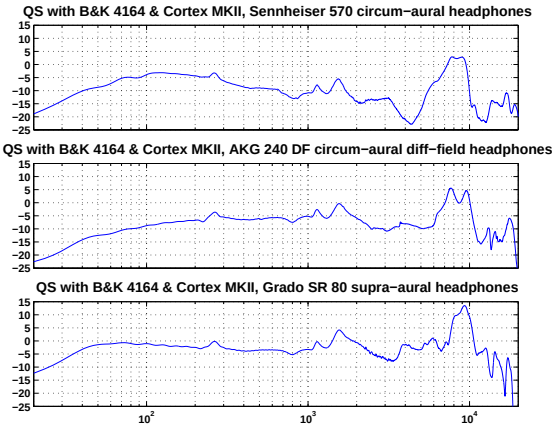

Fig. 10. Different quality headphones measured on the Cortex MK II applying the built-in microphones, no smoothing; a) less expensive Sennheiser 570 headphones, b) middle: less expensive *diffuse-field* equalized AKG 240DF circum-aural headphones, c) below: (morfe expensive) generally well appreciated Grado SR80 supra-aural headphones.

which is similar to response of a diffuse-field equalized headphone, the AKG 240 DF. In practice, there is quite a lot differences between various models, and there is so simple truth what a good headphone frequency response should be.

Judging the UDADU1a tube headphones with the HD 580, the former have less smooth response, but well acceptable considering their very different working principle. The largest difference is between 3-6 kHz where they have an extra 10 dB level decrease, which is highly affected by the standing waves in the ear canal.

Effect of ear canal cross-section?

The antiresonance between 4-5 kHz and a resonance just below 2 kHz was noted to be common to all ear replicas with different lengths, see Figs. 6 and 8. This discovery is noteworthy, since the longitudinal standing waves do not explain this, as in that case the resonances should shift in frequency, but in practice they do not. Even more, the same resonance structure is noted in the headphone measurements (see Figs. 9 and 10) with the Cortex dummy head, hence it is not a question of the tube headphone characteristics. However, with blocked PTFs (see Riederer 1998b), the resonance structure disappears. This means that the phenomenon a property of the ear canal, but not its length, hence it must be its cross-section. It remains to be a puzzle how in a cavity less than 1 cm in diameter resonances in such low frequencies (around 2 and 4 kHz) take place, the presented formulae do not explain this.

CONCLUSIONS

In order to make possible accurate 3-D sound investigations on a 306-channel MEG equipment, special tube headphones were devised (Unides Design Ay.) for the Brain Research Unit at the Low Temperature Laboratory (LTL), HUT. The principal concern was to obtain maximally flat wide-range frequency response (100-10000 Hz), without any active equalization in the device.

The devised UD-ADU1a insert-type tube headphones consist of two discrete, identical units, which are driven with any normal stereophonic amplifier yielding at least 100W RMS sound power. The modified dynamic transducer with passive electronics is enclosed in molded aluminum box (170*120*55 mm^3), to which is fitted a flexible PVC tube (l = 3.00 m, $Ø_1$ = 4 mm, $Ø_2$ = 10 mm). The tube is fitted to a smaller laboratory rubber tube of (l = 0.25 m, $Ø_1$ = 2.5 mm, $Ø_2$ = 5 mm) that is fixed to listeners' ears with replaceable EAR-tips. The earplugs attenuate the background

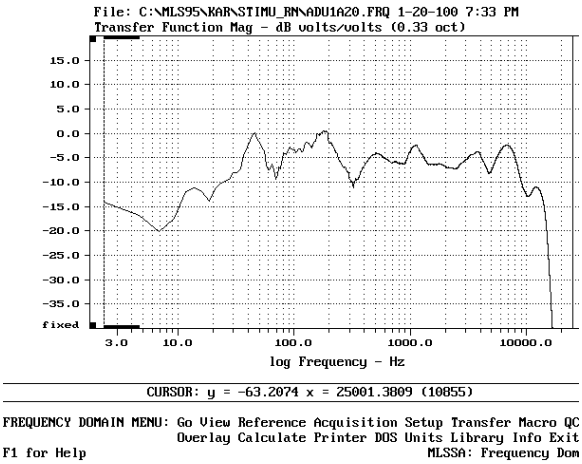

Fig 11. Frequency response of an early version of the UD-ADU1b prototype measured with a Type 2 Artificial Ear and MLSSA-software/hardware (screenshot from the software); smoothed 0.33 octave.

sound typically 15-20 dB (depending on the frequency) and allow a precise positioning of the sound source on the subject.

The problem of doing measurements of sound sources in the human ear canal was discussed in detail in the study. Various measurements were performed applying different models for replicating the average human auditory canal. The frequency responses demonstrate a rough general behavior, with significant variation between the cases. The UD-ADU1a headphones frequency response is 30-13000 Hz (±10 dB) measured with a Type 2 Artificial Ear (ITU-T 1993), 1/3-octave smoothed.

The underlying investigation showed a common characteristic for the human auditory canal offering possibly new information to the field. The collected measurements showed that the ear canal produces a significant antiresonance between 4-5 kHz and a resonance just below 2 kHz *independently from its length*. The resonance structure must be caused by the cross-section of the ear canal, but its grounds remain to be solved. The generally accepted fact has been that only above 5 kHz the shape (i.e., cross-section) of the ear canal becomes significant (e.g., Blauert 1997, p. 302).

Future work

The UD-ADU1a tube headphones outrun the original design criteria and they are used in collaborative neurophysiological research between LCE and LTL and psychophysiological research between LCE and ACO. Forthcoming experiments will address the perceptual differences between 3-D sound reproduction of the tube headphones and high-quality circum-aural headphones. Noteworthy is, that the reproduction position of the tube headphones, i.e., sound source at the blocked ear canal, matches well with the HRTF measurement position (Riederer 1998ab).

The next generation model UD-ADU1b is on an early prototype stage. Fig. 11 shows the frequency response with a minimal ±5 dB deviation in the frequency band of 30-9000 Hz, 0.33 octave smoothed. Currently, there is no need for the upgrade version in the author's research that applies synthetic 3-D sound stimuli including flat equalization of the UD-ADU1a tube headphones.

Final word: what is correct?

Real question is what kind of an ear replica comes closest to the human (subjective) perception? Traditionally a good loudspeaker is considered to have a smoothly declining (e.g., 2-3 dB per octave) power spectrum in the diffuse-field and an exactly flat

frequency response in the free-field – typically measured at a 1.0 m (or 2.5 m) distance at the axis of the upper element ("ear level"). For (circum-aural or supra-aural) headphones the case is less clear because of the close proximity of the sound source to the eardrum, and for insert headphone inside the auditory canal the matter is even much less straightforward. Because of the ear canal occlusion, standing waves cause resonance structures that would not exist without it complicating the situation further at high frequencies, above ca. 4 kHz. The canal dimensions are individual and vary even greatly according to sex and age etc. The standards (e.g., ITU-T P.57) do not consider individuality and offer only an average solution for measuring headphones that might offer a result that is far off from the individual percept.

The reader is welcomed to make an own judgment of the sound quality of the UD-ADU1a tube headphones during the demonstrations at the AES 21st Conference in St. Petersburg, June 1-3, 2002.

ACKNOWLEDGEMENTS

The authors warmly thank the Low Temperature Laboratory (HUT) for lending the Brüel&Kjær Artificial Ear with Ear Canal extension. The first author expresses his humble gratitude for the financial support of the Finnish Graduate School of Electronics, Telecommunication and Automation, supported by the Finnish Academy / Ministry of Education / State of Finland.

REFERENCES

Blauert J., 1997. *Spatial Hearing. The Psychophysics of Human Sound Localization.* Revised Edition, MIT Press, Cambridge, Massachusetts, USA, 494 p.

Brüel&Kjær, 1997. *Sound & Vibration Catalogue*. Brüel&Kjær, Naerum, Denmark, 151 p.

Hammershøi D. and Møller H., 1996. Sound transmission to and within the human ear canal. *J. Acoust. Soc. Am.*, vol. 100, no. 1, pp. 408-427.

IEC 318, Electroacoustics - Simulators of human head ear - Part 1: Ear simulator for calibration of supraaural earphones. *International Electrotechnical Commission* (IEC), International standard IEC 60318-1 Ed. 1.OB, 7/1998.

IEC 711, Occluded-ear simulator for the measurement of earphones coupled to the ear by ear inserts. *International Electrotechnical Commission* (IEC), International standard IEC 60711 Ed. 1.OB.

IEC 959, Provisional head and torso simulator for acoustic measurement on air conduction hearing aids. *International Electrotechnical Commission* (IEC), Technical report IEC 959, 1990.

IEC, 1990. Provisional head and torso simulator for acoustic measurements on air conduction hearing aids. *International Electrotechnical Commission* (*IEC*), international standard IEC/TR0 60959 (1990-05). Ed. 1.0. ICS code: 17.140.50 Electroacoustics.

ISO 868, Plastics and ebonite - Determination of indentation hardness by means of a Durometer (Shore hardness). *International Standardization Organization*, International Standard 868, 1985.

ITU-T, 1993. Telephone transmission quality objective measuring apparatus. Artificial Ears. *Telecommunication standardization sector of International Telecommunication Union.* Recommendation P.57, 3/1993, Helsinki.

J. O. Smith and J. Abel, 1995. The Bark bilinear transform. *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, New York, October 15-18, 1995.

M. B. Rappaport and H. B. Spraque, 1952. The effects of improper fitting of stethoscope to ears on auscultatory efficiency, *American Heart Journal*, Jan. 1951, p. 713-715.

M. Karjalainen, 1990. DSP software integration by object-oriented programming: A case study of QuickSig, *IEEE ASSP Magazine*, Apr 1990, pp. 21–31.

Mairas A., Järvinen A. and Piirilä E., 1999. Using tubephones for wideband auditory stimulation in brains research. Proc. of Forum Acusticum, Berlin, 1999.

Møller H., Hammershøi D., Jensen. C. B. and Sørensen M. F., 1995. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, vol. 43, no. 4, pp. 203–217.

Poldy, C.A. *Headphones*. Loudspeaker and headphone handbook. Focal press 1994, pp.493-577

Riederer K. A. J and Backman J., 1998. *Frequency response of stethoscopes.* Proceedings of the Nordic Acoustical Meeting 1998 (NAM98), Stockholm, Sweden, 6-9.9.1998.

Riederer K. A. J and Karjalainen M., 1998. DSP aspects of head-related transfer function measurements. In *IEEE Nordic Signal Processing Symposium, NORSIG'98*, Vigsø Holiday Resort, Denmark, 8-11.6.1998.

Riederer K. A. J, 1998a. *Head-related transfer function measurements*. Master's Thesis. Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, 134 p.

Riederer K. A. J, 1998b. Repeatability analysis of head-related transfer function measurements, *105th Audio Engineering Society Convention*, San Francisco, Sept. 26-29. Preprint no. 4846, 62 p.

Rife D. D., 1996. *MLSSA Reference Manual Version 10.0*. DRA Laboratories. USA.

Rossing, T.D. The science of sound. Second ed. Addison-Wesley publishing Company Inc. 1990 USA. pp.686.

S. Uosukainen, 2002. *Personal communication* 17.4.2002, VTT Technical Research Centre of Finland, VTT Building and Transport, Espoo, Finland.

Sennheiser electronic KG, 1996. Electret microphone capsule KE 4-211-2 industry information. Germany.

Shaw E. A. G. and Teranishi R., 1968. Sound pressure generated in an external ear replica and real human ears by a nearby sound source. *J. Acoust. Soc. Amer*., vol. 44, pp. 240-249.

Unides Design, 1999. *UDMPA10d Dual Channel Microphone Preamplifier Data 02.1999*. Unides Design, Avoin yhtiö, Helsinki, Finland, 4 p.

Wightman, F. L., & Kistler, D. J., 1989. Headphone simulation of free-field listening I: stimulus synthesis. *J. Acoust. Soc. Amer*., vol. 85, pp. 858-867.