Head-related transfer functions of human subjects

Møller, Henrik; Sørensen, Michael Friis; Hammershøi, Dorte; Jensen, Clemen Boje

Published in:
Journal of the Audio Engineering Society

Publication date:
1995

Head-Related Transfer Functions of Human Subjects*

HENRIK MØLLER, *AES Member,* **MICHAEL FRIIS SØRENSEN,** *AES Member,*
DORTE HAMMERSHØI, *AES Member,* **AND CLEMEN BOJE JENSEN,** *AES Member*

Acoustics Laboratory, Aalborg University, Aalborg, Denmark

Head-related transfer functions (HRTFs) were measured on 40 human subjects for 97 directions of sound incidence, covering the entire sphere. Individual HRTF data for the median, horizontal, and frontal planes are presented in the frequency domain. Measurements were made synchronously at both ears, thus making the time representations, that is, the head-related impulse responses (HRIRs), valid also when interaural time differences are considered. The measurements were made at the entrance to the blocked ear canal. Sound at this point contains full spatial information, and the interindividual variation is lower than at the open ear canal.

0 INTRODUCTION

A head-related transfer function (HRTF) is a transfer function that, for a certain angle of incidence, describes the sound transmission from a free field to a point in the ear canal of a human subject. Knowledge of human HRTFs is essential in the design and evaluation of artificial heads and in computer synthesis of binaural signals, such as in virtual reality applications or for auralization in room modeling systems. HRTFs can also be used to compute diffuse-field to free-field corrections for hearing thresholds and equal-loudness contours, and to determine design goals for headphones. The interest in the present investigation, though, is use in the binaural technique. A preliminary report of the investigation was given at the 92nd Convention of the Audio Engineering Society in Vienna, Austria, in 1992 [1].

0.1 Binaural Technique

The idea behind the binaural technique is the following. The input to the hearing consists of two signals—the sound pressures at each of the eardrums. If these are recorded in the ears of a listener and reproduced exactly as they were (usually through headphones), then the complete auditory experience is assumed to be replicated, including timbre and spatial aspects.

The reason why the listener is able to perceive the direction to a sound source is that the sound on its way to the two ears is exposed to filtering corresponding to the particular direction. The filtering is due to reflections and diffractions from the human torso, head, and pinna, and is described by the HRTFs. The hearing is able to "recognize" the filtering and thus to determine the direction to the source. In most cases the amplitude versus frequency responses as well as the arrival times are different in the two ears. A thorough description of the directional cues used by the hearing is given by Blauert [2], [3].

A system that implements the binaural technique is called a binaural system. In most systems the listener is replaced by an artificial head during the recording. The artificial head should provide the same directional filtering, that is, have the same HRTFs, as the listener. Thus knowledge about human HRTFs plays an important role in the design and evaluation of artificial heads. A review of the binaural technology is given by Møller [4].

Binaural signals need not originate in a recording with a physical head. It is possible to synthesize binaural signals on a computer and simulate a sound being played in a room and recorded binaurally. Apart from simulating the room transmission, the computer must filter with the relevant HRTFs. The computer program should therefore include a database of HRTFs for a large number of directions. Since the filtering is normally carried out as convolutions, HRTFs for this purpose are most conveniently given in the time domain, that is, as head-related impulse responses (HRIRs).

Because of anatomical differences between people, there are individual differences in human HRTFs. In most applications, however, all listeners must listen to the same binaural signals. The artificial head—and in computer synthesis of binaural signals, the HRTF data-

* Manuscript received 1994 July 11; revised 1995 January 9.

base—should therefore represent a typical human head. Knowledge about HRTFs for a large number of human subjects is needed.

If the spread between HRTFs for human subjects is large, it cannot be expected that the binaural technique will work for an entire population. Information about the spread is therefore important in an evaluation of the objective basis for the binaural technique.

0.2 Previous Measurements of HRTFs

Measurements of HRTFs of human subjects have already been described in the literature [2], [3], [5]–[27]. The investigations aim at different goals, which explains why the measuring technique, the number of source directions, and other parameters vary considerably. For instance, some measurements were based on sinusoidal tones [5]–[7], [10]–[12], [18], some on one-third-octave noise [9], [27], some on various impulse techniques [13], [15]–[17], [21], [22], and some on various noise or pseudorandom noise techniques [19], [20], [23], [24], [26].

The point in the ear canal where the measurement is made is called the reference point. Also the choice of this varies. Robinson and Whittle [8] measured 6–9 mm outside the ear canal entrance, Wiener [6], Shaw [10], Burkhard and Sachs [12], Morimoto and Ando [16], and Okabe and Miura [24] chose a point at the entrance to the ear canal, Mehrgardt and Mellert [13] a point 2 mm from the entrance inside the ear canal, Platte and Laws [14], Platte [15], and Genuit [17] 4 mm from the entrance, Schmitz and Vorländer [23] 4–5 mm from the entrance, Blauert [2], [3] 5 mm from the entrance, Pösselt et al. [18] also 5 mm from the entrance, but blocking the ear canal at this point, Middlebrooks et al. [21] and Middlebrooks and Green [22] at least 5 mm from the entrance, Middlebrooks [26] 5–10 mm from the entrance, Searle et al. [11] 10 mm from the entrance, Hellstrom and Axelsson [27] 1–3 mm from the eardrum, Wenzel et al. [19] and Wightman and Kistler [20] 1–2 mm from the eardrum, and Jahn [9] at the eardrum. Wiener and Ross [5] measured 1 mm from the eardum and also at the entrance to the open and the blocked ear canal, Yamaguchi and Sushi [7] at the eardrum and at the entrance to the open ear canal.

The sound pressure changes along the ear canal, and consequently the results, cannot be compared directly. Also the presence of a microphone in the ear canal may affect the measurement. The change in sound pressure along the ear canal has been shown experimentally [5], [7], [13], [21], [27] and by Djupesland and Zwislocki [28] and Hammershøi and Møller [29].

The results of previous experiments are most often presented as amplitude responses in the frequency domain. Even when time information is available, a sufficiently precise definition of the time axis may be lacking if measurements are not carried out simultaneously at the two ears. In some cases the number of sound directions is low, and often no directions below the horizontal plane are included. The number of subjects is often insufficient for an evaluation of interindividual differences.

Especially the variation in reference point and the differences in microphone technique are considered major impediments for a comparison of results. In the present investigation the entrance to the blocked ear canal has been chosen as the reference point, since it has a number of advantages, as will be seen.

0.3 Model of Sound Transmission

A model of the transmission from a free sound field to the ear canal is described by Møller [4] and verified by Hammershøi and Møller [29]. The sound transmission is divided into two parts, a direction-dependent part that creates all directional cues, and a part that is independent of direction.

The *direction-dependent* part of the sound transmission is outside the ear canal. It consists of the sound transmission from the free-field sound pressure P_1 at the head center position, but with the listener absent, to the Thevenin sound pressure P_2 at the entrance to the ear canal. P_2 does not exist physically in a listening situation, but for measurement purposes it can be found at the entrance to the ear canal when the ear canal is blocked, as, for instance, by an earplug. The direction-dependent part is thus described by the transfer function

$$\frac{P_2}{P_1}(\phi, \theta) = \frac{\text{sound pressure at entrance to blocked ear canal}}{\text{sound pressure at center position of head}} \tag{1}$$

where ϕ and θ are azimuth and elevation, respectively.

The *direction-independent* part of the model is given in Fig. 1. Fig. 1(a) is a sketch of the physics of the human external ear, and Fig. 1(b) is the corresponding analog model.

Everything outside the ear canal is modeled by a The-

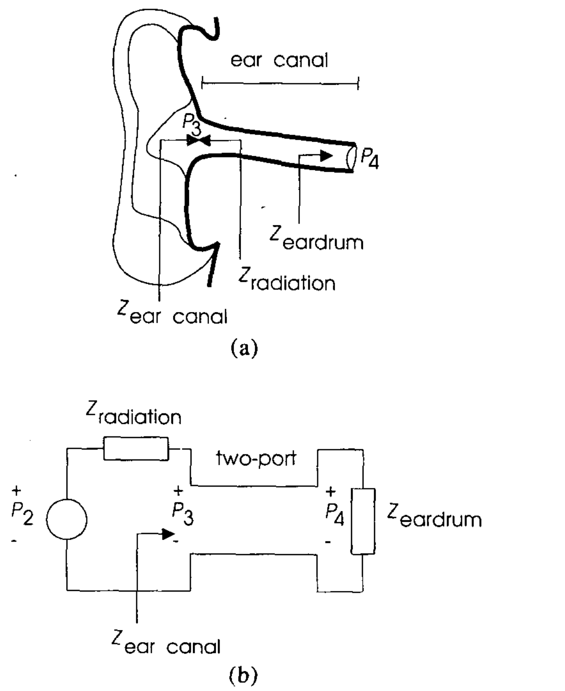

Fig. 1. Sound transmission through external ear. (a) Sketch of anatomy. (b) Analog model. (From Møller [4].)

venin equivalent, with the generator sound pressure P_2 and the generator impedance $Z_{radiation}$. The sound pressure at the input to the ear canal is denoted by P_3. In the frequency range of interest, the canal acts as an acoustic two-port terminated by the eardrum impedance $Z_{eardrum}$, and the resulting pressure at the eardrum is denoted by P_4.

The direction-independent transmission consists of the pressure division at the entrance to the ear canal, given by

$$\frac{P_3}{P_2} = \frac{Z_{ear\ canal}}{Z_{ear\ canal} + Z_{radiation}}$$

and the transmission along the ear canal P_4/P_3.

Obviously all three sound pressures P_2, P_3, and P_4 contain the spatial information, and they can all be used for the measurement of HRTFs and for binaural recordings. (The same applies to all sound pressures along the ear canal.) P_2 is the most convenient to use, as it does not require the insertion of microphones deeply into the ear canals, and the microphones can be placed in an earplug without interfering with the measurement. Furthermore, sound pressures at other points are unnecessarily influenced by individual differences in the ear canals, a matter that is further discussed by Hammershøi and Møller [29]. P_2 can be characterized as the sound pressure that includes the directional information and, in addition to that, the minimum amount of individual information.

Based on these considerations a reference point outside the blocked ear canal has been chosen in the present investigation. If nothing else is specified, the term HRTF will, for the remainder of this paper, refer to the head-related transfer function measured at the blocked ear canal, as given by Eq. (1). The term HRIR will refer to the time-domain representation of the same transfer function.

It was shown previously [4] that headphones for the binaural technique should have (or be equalized to) a flat frequency response when measured at the point in the ear canal where the recording is made. However, if the recording is made outside a blocked ear canal, an extra correction is needed, unless the reproduction involves a pressure division equal to the one that would have been during the recording [Eq. (2)]. The reproduction side of the binaural technique is evaluated in a parallel investigation at our laboratory [30], and in order to provide data for that, measurements of the pressure division P_3/P_2 were included in the present investigation.

0.4 Goal of Present Investigation

It is the purpose of the present investigation to provide knowledge about the HRTFs and the corresponding HRIRs for a large number of people and for a large number of directions, covering the entire sphere.

In addition, knowledge should be obtained about the pressure division P_3/P_2. As it is assumed independent of direction, it only needs to be measured for one direction. However, with the purpose of confirming the directional independence it will be measured for a few directions of sound incidence.

1 METHOD

Measurements were carried out on subjects standing in an anechoic chamber. Eight loudspeakers were placed in an arc and the subjects were rotated to obtain measurements from 97 directions, covering the entire sphere. Impulse responses were measured for the transmission from the voltage at the input of the power amplifier to the output of the measuring microphone, placed to measure P_1, P_2, or P_3. The impulse responses were measured using the maximum-length-sequence (MLS) technique, and sample synchronous measurements were made at the two ears. HRTFs and pressure divisions P_3/P_2 were obtained through Fourier transformation of the measured impulse responses, followed by appropriate divisions. Subsequent inverse Fourier transformations of the HRTFs produced HRIRs.

For each subject 107 measurements were made. P_2 was measured for 97 directions with miniature electret microphones. Both P_2 and P_3 were measured for five directions with probe microphones (with the incidence of the sound wave being front, back, left, right, and above). For each loudspeaker P_1 was measured with a standard laboratory microphone.

1.1 Subjects

40 subjects participated, 22 males and 18 females; 31 were randomly chosen students and 9 were staff members. The age range was 21–40 years. All subjects had normal hearing, and none had reported ear abnormalities that might affect the middle ear function. Photos of the subjects were made, enabling future work on relating the subjects' anatomy to the HRTFs. The same subjects participated in the investigation of headphone transfer characteristics [30].

1.2 Microphones

The choice of microphones for measurements in the human external ear is a compromise between satisfying sensitivity and frequency response, and acceptable dimensions. For the measurement of the Thevenin pressure P_2, a Sennheiser KE 4-211-2 miniature microphone was used. This is an electret microphone, cylindrical in shape, with the dimensions 4.75 mm (diameter) by 4.20 mm (height). Two selected microphones were used throughout the experiment, one for each ear.

The microphones were mounted in EAR earplugs placed in the ear canals. The microphone was inserted in a hole in the earplug, and then the soft material of the earplug was compressed during insertion in the ear canal. As the earplug relaxed, the outer end of the ear canal was completely filled out. The end of the earplug and the microphone were mounted flush with the ear canal entrance. The placement of the microphone is shown in Fig. 2.

For the measurement of P_3 it was reckoned that the miniature microphone was so big that it would disturb the sound field, and a probe microphone was therefore used. This type of microphone would not disturb the sound field in the ear canal, since the interference of the probe tip is minimal, but it has the disadvantage of a low sensitivity and a nonflat frequency response. The probe microphone was a Brüel & Kjær 4182 with a 45-mm metal tube extended by 5 mm of flexible plastic tube (outer diameter 1.65 mm). The acoustic impedance of the probe was approximately 10^9 N · s/m^5, which at 500 Hz corresponds to the impedance of a 0.05-cm^3 volume. At higher frequencies the corresponding volume decreases, and loading of the ear canal by the tube was therefore considered negligible in the frequency range of interest.

During measurements the probe microphone was attached to the external ear by a metal strap. The strap and the tube were adjusted individually to fit the shape of the ear. To avoid displacement of the probe microphone during the experiment, it was fixed along the subject's neck with surgical tape, and the position of the tip was controlled before and after each single measurement.

Even though efforts were undertaken to keep the housing of the probe microphone close to the subject's neck, some disturbance of the sound field could be expected from the microphone housing. Therefore P_2 measurements were also carried out with the probe microphone for the same directions as P_3. In this way the microphone housing had the same influence on P_2 and P_3, and the influence on the pressure division was eliminated. Due to the influence of the microphone housing, the measurements with the probe microphones are considered less suitable for determining HRTFs and HRIRs than measurements with miniature microphones.

In order to minimize the error due to a displacement of the probe tip between measurements of P_2 and P_3, the probe measurements were carried out in the following way. P_2 was measured first. Then the earplug was carefully removed with as little disturbance of the probe microphone as possible. Immediately afterward P_3 was measured. Fig. 3 shows the probe microphone attached to a subject's ear and neck. The pressure frequency responses for the miniature and the probe microphones are given in Fig. 4.

The reference sound pressure P_1 from each loudspeaker was measured with a 6.4-mm (1/4-in) microphone (Brüel & Kjær 4136). The microphone was placed at the position where the middle of the subject's head would be during measurement and with an orientation giving 90°

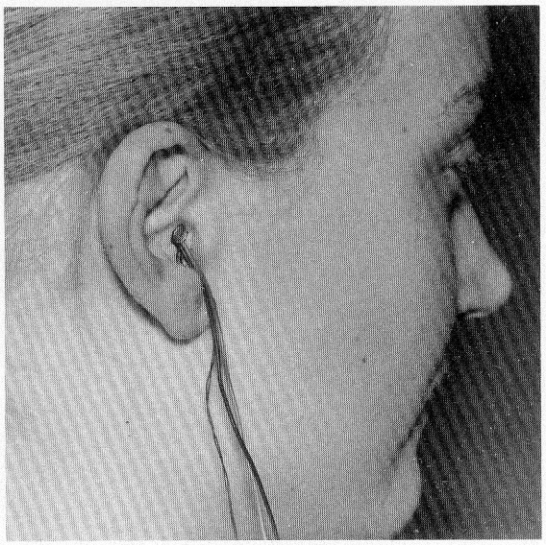

Fig. 2. Miniature microphone mounted in ear of subject to measure blocked ear canal pressure P_2.

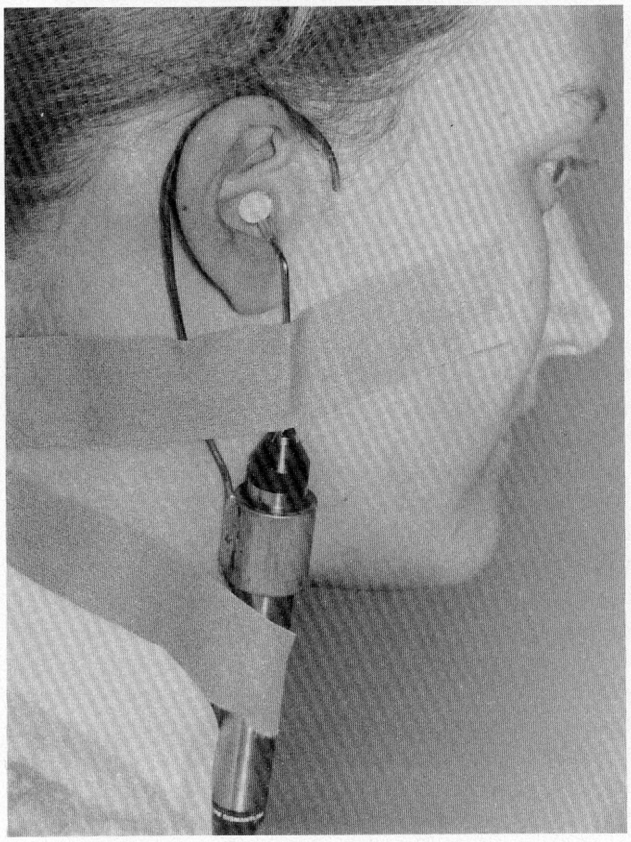

Fig. 3. Probe microphone mounted on human subject to measure blocked ear canal pressure P_2.

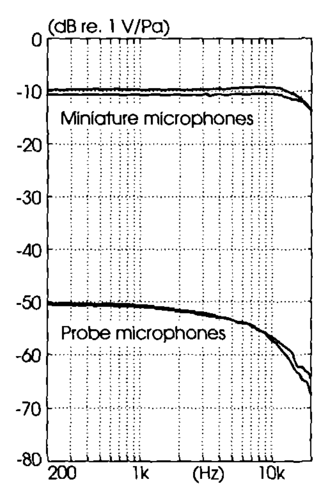

Fig. 4. Pressure frequency responses of the two miniature microphones (Sennheiser KE 4-211-2), including 20-dB front-end amplifiers, and the two probe microphones (Brüel & Kjær 4182), including 45-mm metal tube and 5-mm plastic tube.

incidence of the sound wave for all loudspeakers. In this way the P_1 measurements were influenced minimally by the presence of the microphone in the sound field, and equally influenced for all loudspeakers. In order to reduce the difference between the free-field and the pressure responses of the microphone, it was used without protection grid.

The sensitivities at 1 kHz of all microphones were measured every working day during the two months' measurement period. Although deviations from the nominal values were small (below ± 0.3 dB), appropriate corrections were made in the final data processing.

1.3 Free-Field Setup

The measurements were carried out in an anechoic chamber with a free space between the wedges of 6.2 m (length) by 5.0 m (width) by 5.8 m (height). Measurements were made with sound arriving from a total of 97 directions, covering the entire sphere. The convention for indicating the direction is as follows: ϕ defines azimuth and θ elevation; $(\phi, \theta) = (0°, 0°)$ is the direction in front of the subject. Positive values of ϕ indicate directions to the left of the subject, and positive values of θ indicate directions above the horizontal plane. Thus the entire sphere is covered for $-180° < \phi \leq 180°$ and $-90° < \theta \leq 90°$.

Both ϕ and θ were chosen in steps of 22.5°. An exception was at $\theta = \pm 67.5°$, where the step size for ϕ was chosen to be 45°, since 22.5° would cause the directions to lie rather close on the sphere. The distance from the sound source to the center of the subject's head was chosen to be 2 m, except for $\theta = -67.5°$, where the existence of a net floor made a reduction to 1.95 m necessary. No measurements were made at $\theta = -90°$, that is, from a position directly below the subject.

Eight loudspeakers were fixed in an arc of a circle with θ ranging from $-67.5°$ to 90° in steps of 22.5°. The setup in the anechoic chamber is seen in Fig. 5. The subject stood on a rotatable platform, which could be fixed at the desired values of ϕ. The subject was standing on the platform in a natural upright position, and a small backrest mounted on the platform helped the subject to stand still. The position of the platform and an individual adjustment of its height caused an imaginary point right between the subject's ears to be in the center of the sphere.

For help in the control of the horizontal position and the orientation of the subject's head, the subject had a paper marker on top of his or her head. This marker was observed through a video camera placed in the direction $\theta \approx 90°$ and shown on a movable monitor in front of the subject. Using this, the subject could correct position and azimuth (Fig. 6).

The operators had a similar monitoring system for observing the subject's exact position and for verifying that the subject did not move during each single measurement. If movements were observed, the result was discarded and the measurement repeated.

The loudspeakers were 70-mm diaphragm midrange units (Vifa M10MD-39) mounted in 155-mm-diameter hard plastic balls. The free-field response of a loudspeaker is shown in the frequency and time domains in Fig. 7.

1.4 Measuring Setup

The general-purpose measuring system known as MLSSA was used (*m*aximum-*l*ength-*s*equence *s*ystem *a*nalyzer, DRA Laboratories). The MLS method offers a number of advantages compared to traditional frequency and time domain techniques. It is basically noise immune, and combined with averaging, the achieved sig-

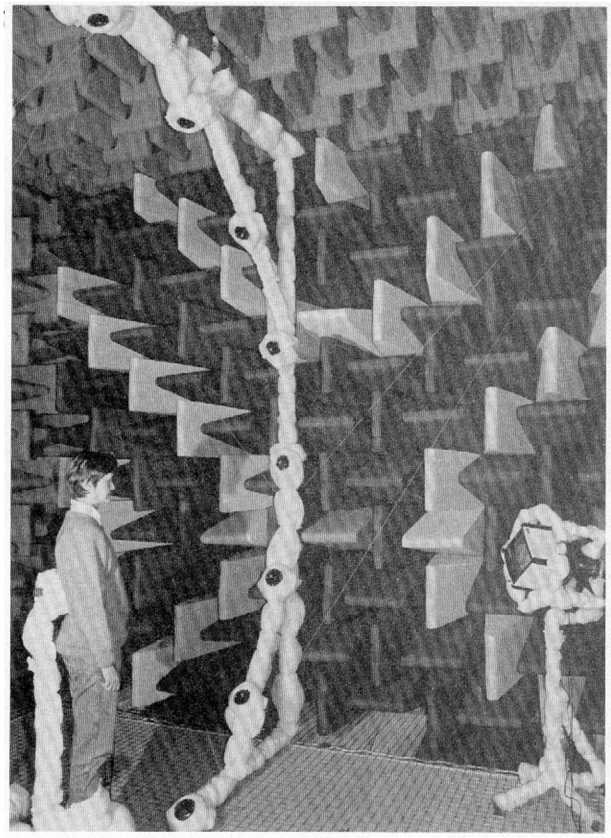

Fig. 5. Free-field setup in anechoic chamber.

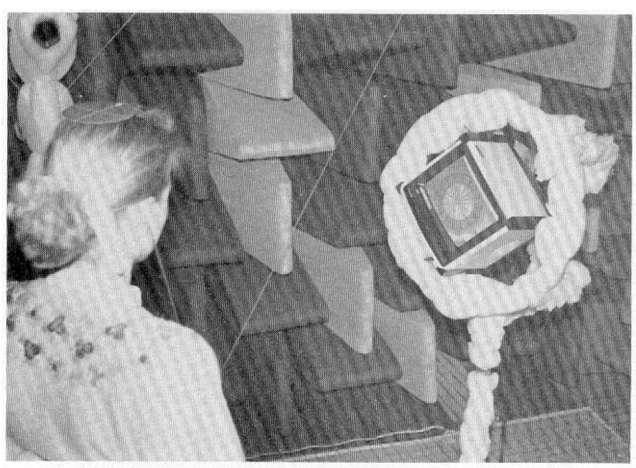

Fig. 6. Subject with direction marker. Note that for phototechnical reasons monitor is closer to subject than during measurements.

nal-to-noise ratio is high. A thorough review of the MLS method is given by Rife and Vanderkooy [31].

For the purpose of measuring at both ears simultaneously, two MLSSA systems were used, coupled in a master–slave configuration by a custom-made synchronization unit allowing sample synchronous measurements. The sampling frequency of 48 kHz was provided by an external clock.

The 4-V peak-to-peak stimulus signal from the master MLSSA board was sent to a power amplifier (Pioneer A-616), which was modified to have a calibrated gain of 1. The output was directed through a switch box to the loudspeaker in the measurement direction. The free-field sound had a level of 75 dB(A) at the subject's position, a level where the stapedius was assumed to be relaxed.

The microphone (miniature as well as probe) signals were sent through measuring amplifiers (Brüel & Kjær 2607). To avoid frequency aliasing, the 20-kHz Chebyshev low-pass filters of the MLSSA boards and the 22.5-kHz low-pass filters of the measuring amplifiers were used. Also the 22.5-Hz high-pass filters on the measuring amplifiers were active.

Preliminary measurements, using an MLS length of 65 535 points, showed that the minimum length offered by MLSSA (4095 points) was sufficient to avoid time aliasing. In order to achieve a high signal-to-noise ratio, the recording was averaged over 16 times, called pre-averaging in the MLSSA system. Even with this averaging the total time for a measurement was as short as 1.45 s. During this period the subjects were normally able to stand still.

1.5 Data Processing

Results of the measurements were impulse responses for the transmission from input to the power amplifier to output of the measuring amplifier. In order to obtain the wanted information, some postprocessing was needed. This was carried out using the program MATLAB (The MathWorks, Inc.).

The measured impulse responses included an initial delay, corresponding to the propagation time from the loudspeaker to the measuring point (approximately 6 ms). All responses were very short with a duration of only a few milliseconds. Therefore only samples from 256 through 511 were processed (time from 5.33 to 10.65 ms). The restriction to this time window eliminated reflections from the monitor in the anechoic chamber.

For determination of the HRTF, the impulse responses to P_1 and P_2 were Fourier transformed, and a complex division was carried out in the frequency domain. As the same electric equipment was involved during the measurement of P_1 and P_2, its effects cancel. Since the miniature microphone was used for the measurement of P_2 and the Brüel & Kjær 4136 for the measurement of P_1, the result of the calculation included the ratio between the pressure response of the miniature microphone and the 90° free-field response of the 4136 microphone.

The ratio between the two microphones' pressure responses was found by measuring the same sound field with the two microphones very close together. This ratio was then removed by a complex division in the frequency domain. At last the result only included the ratio between the 90° free-field response and the pressure response for the 4136 microphone. This ratio was found in the data sheet to have its maximum value of 0.6 dB at 20 kHz, and nothing was done to correct for this.

The HRIR was determined through an inverse Fourier transform of P_2/P_1. Before the transformation, P_2/P_1 was low-pass filtered at 20 kHz. The pressure division P_3/P_2 was determined with a procedure similar to that for P_2/P_1, and since the probe microphone was used for the measurement of both P_2 and P_3, the pressure response of the probe microphone cancels.

Measurements of P_1 and of the frequency responses of the microphones were carried out repeatedly throughout the experimental period. Differences were small, though, and the same set of responses was used in the processing of all data.

1.6 Sources of Errors

The most obvious sources of errors are disturbing reflections, electrical and acoustical noise, and improper placing of the microphones and the subjects. In the free-field setup the only objects that could give rise to reflections within the time window were other loudspeakers and the arc holding the loudspeakers, the backrest, and the platform on which the subject was standing. Reflections from the loudspeakers and the arc were eliminated by means of sound-absorbing material wrapped around the loudspeakers and the arc. The effect of the coverage was verified through free-field measurements at the subject's position, with parts of the loudspeaker arc re-

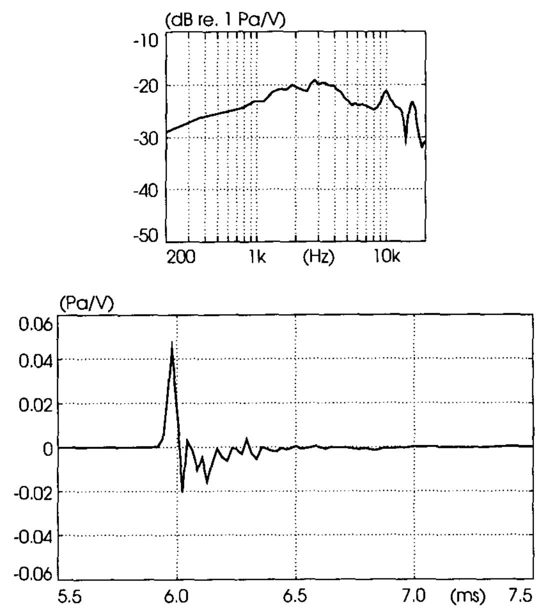

Fig. 7. Typical frequency response and impulse response for loudspeakers (Vifa M10MD-39 mounted in 15.5-mm hard plastic ball), measured at a distance of 2 m with Brüel & Kjær 4136 microphone, 90° incidence.

moved and repositioned.

The influence of reflections from the backrest was eliminated by covering it with sound-absorbing material. The effect was verified through measurements on a head and torso simulator (Brüel & Kjær 4128), measured in expected worst-case directions, with and without backrest.

Possible reflections from the platform would only occur in the time window for elevations between $-67.5°$ and $-22.5°$. The platform was also wrapped with sound-absorbing material, and the foot plate was shielded in the direction of the loudspeakers with a sound-absorbing fiber mat. For practical reasons the effect of the coverage of the platform was not verified through measurements.

The signal-to-noise ratio cannot be expressed as a single figure, since it depends on the response being measured. For illustration, the total noise level was found by repeating a measurement, but with the loudspeaker electrically replaced by a resistor. All gain settings were as in the original measurement. Examples with the two types of microphones are given in Fig. 8. The signal-to-noise ratio for the miniature microphones is typically around 70 dB and only below 50 dB at frequencies where the transfer functions exhibit very low values. For the probe microphones the corresponding figures are approximately 15 dB lower.

As the subject during the measurements had a natural upright position without fixation of the head, errors due to wrong position and orientation could not be avoided, even when the position marker was placed carefully. The error due to an incorrect *position* of the subject is considered in Section 2.1 in connection with a presentation of the HRIRs from all subjects from one direction. Errors due to an incorrect *orientation* of the subject's head are evaluated in connection with the presentation of interaural time differences in Section 2.3.

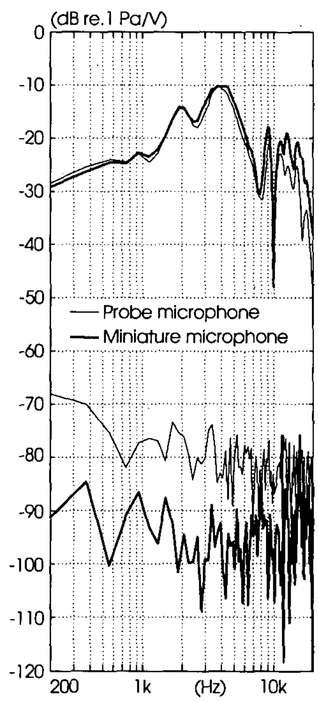

Fig. 8. Illustration of signal-to-noise ratio. Upper curves—typical measurements; lower curves—corresponding measurements with all settings unchanged, but with loudspeaker replaced by 8-Ω resistor. Data are original measurements, not computed HRTFs.

The total error caused by the above-mentioned factors and by inaccurate microphone placement is evaluated through a series of repetition measurements presented in Section 2.5. That section also contains a comparison of results obtained with the two microphone techniques, the miniature microphone and the probe microphone.

2 RESULTS

A very large amount of data was collected in this investigation, and it is impossible to present everything. The following gives an overview of the data we have chosen to present.

Section 2.1 concentrates on results from one single direction. Starting with the data from one subject, HRTFs and HRIRs are shown and described. Following that, data for all 40 subjects are presented, illustrating interindividual variations in the frequency and time domains.

Results from all directions in the median plane, the horizontal plane, and the frontal plane, covering 41 directions of sound incidence, are presented in Section 2.2. Since HRTFs are most often presented as plots of amplitude versus frequency, only this form is given. Results from the left ear of all subjects are shown.

An important information, which is lost when only amplitude responses are shown, is the interaural time difference (ITD). For the same three planes, this is presented in Section 2.3.

In Section 2.4 the directional independence of the pressure division P_3/P_2 is illustrated with examples from a few subjects. The variation between subjects is also illustrated, but no further analysis is made, since the measurements were mainly made for the benefit of another investigation [30].

Based on a series of repetition measurements, Section 2.5 contains an assessment of the total errors, including inaccurate orientation of the subjects and inaccurate microphone placement.

2.1 Results from a Single Direction

As a simple start of the data presentation, Fig. 9 shows an example of HRTFs and HRIRs for the two ears of one subject. The sound comes from the left side in the horizontal plane, as indicated by the arrow pointing toward the little head in the figure $[(\phi, \theta) = (90°, 0°)]$.

The HRTFs in the upper part of Fig. 9 show that the signal at the right ear is attenuated compared to that at the left ear. At most frequencies the transmission to the left ear exceeds 0 dB, which indicates a pressure buildup due to reflections from the head and pinna. The maximum gain of more than 15 dB is obtained at approximately 6 kHz. Characteristic peaks and dips are seen, mainly above 1 kHz. The transmission to the right ear has an overall shape of a low-pass filter with a cut-off frequency around 4 kHz due to the shadowing effect of the head.

In general the HRTFs converge toward 0 dB at low frequencies, where the presence of the human body does not disturb the sound field. However, still at 200 Hz a gain of 2 dB is seen at the left ear. Of these, approximately 0.4 dB can be explained by the fact that the left ear is some 80 mm closer to the sound source than the middle of the subject's head.

The lower part of Fig. 9 shows the HRIRs. The level difference between the ears is clearly seen. It is also seen that the signal at the left ear arrives slightly before zero, obviously because this ear is closer to the sound source than the middle of the head. The right ear signal arrives later than zero, approximately 600 μs after the signal to the left ear.

Fig. 10 illustrates the same situation, but with curves for all 40 subjects. Unmistakable interindividual variation is seen in the HRTFs, but it is still possible to find a common structure similar to the curves in Fig. 9. This is especially true for the left ear, whereas it is a little more doubtful for the right ear.

The structures of the HRIRs are more difficult to recognize in Fig. 10. A substantial amount of the variation can be explained by inaccurate positioning of the subjects. A displacement of the subject of only 10 mm corresponds to a displacement in time of the HRIRs of about 30 μs, which is clearly visual on the graph.

In Fig. 11 the individual HRIRs have been moved along the time axis to give the best possible fit at the onset of each HRIR, defined as the first sample, where it exceeds 10% of its maximum value. In the upper part of the figure each subject's left- and right-ear HRIRs are shifted by the same amount, while in the lower part the left- and right-ear HRIRs have been allowed to shift independently.

It is now possible to see a clear common structure of the HRIRs, especially in the lower part of Fig. 11. Note that the procedure for the lower part eliminates variations due not only to inaccurate positioning of the subjects, but also to variations in the interaural time difference (ITD).

The amount of time displacement needed to obtain the upper part of Fig. 11 allows an estimation of the inaccuracy in the subject's position. In most cases a displacement of less than 20 μs was used, corresponding to a horizontal sideward position error of less than 7 mm. The same procedure used for HRIRs from the front shows that in most cases the horizontal front–back position error was below 12 mm.

When used for the direction above the subject, the procedure indicates that most errors in the vertical position are below 8 mm. This corresponds well with the step size in the height adjustment of the setup, which was 11 mm.

It should be stressed that since measurements were made synchronously at the two ears, the time information of the HRIRs is not invalidated by the position inaccuracies. The shape of each individual HRIR is still valid, and so is the difference between the arrival times

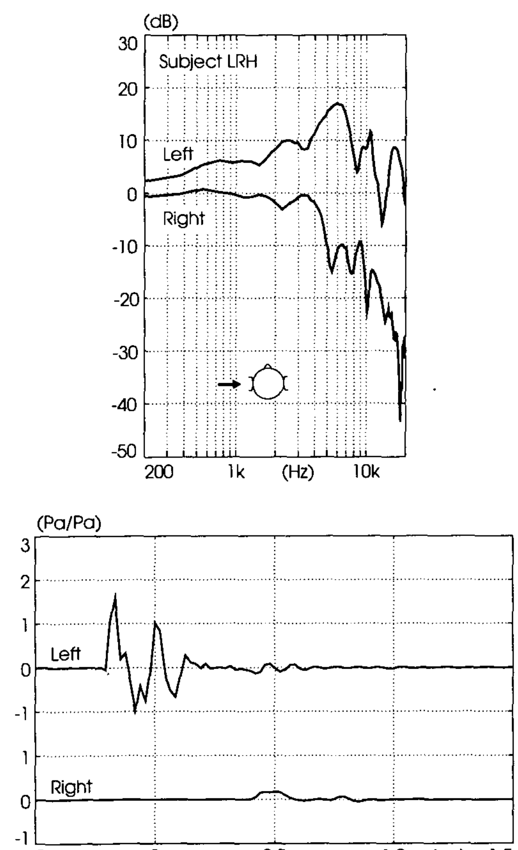

Fig. 9. Example of HRTFs and HRIRs for one subject, direction (ϕ, θ) = (90°, 0°).

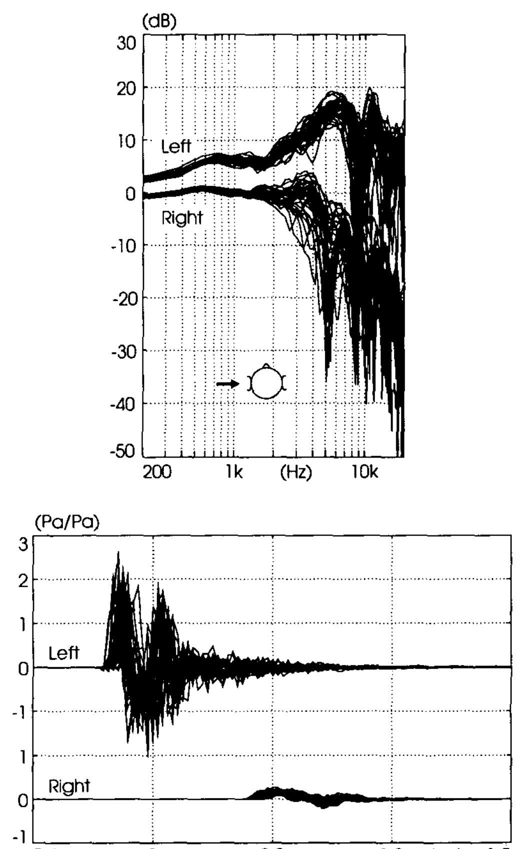

Fig. 10. HRTFs and HRIRs for direction (ϕ, θ) = (90°, 0°), same as in Fig. 9, but for all 40 subjects.

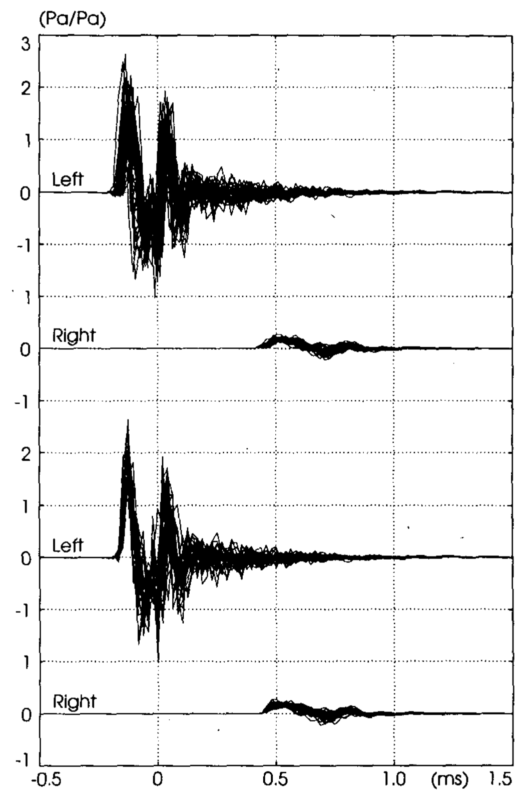

Fig. 11. HRIRs for left ear of all subjects for direction (ϕ, θ) = (90°, 0°). Data are similar to those shown in Fig. 10, but HRIRs are offset in time so they all start at mean arrival time. See text for difference between upper and lower sets of curves.

for the two ears. With a 2-m distance to the sound source, the angle error caused by the inaccuracy in the subject's position mentioned can be neglected.

The importance of the phase of the HRTFs has often been a matter of discussion. A more detailed analysis of some selected HRIRs showed that even when the arrival time has been accounted for, some of them do not represent minimum-phase systems. An excess phase corresponding to 1–3 complex conjugate all-pass sections seems to be common. Examples of HRIRs, their minimum-phase counterparts, and their all-pass phases are shown in Fig. 12. Minimum-phase representations of HRIRs were included in the listening experiments reported in [32], [33].

2.2 Head-Related Transfer Functions for Three Planes

Figs. 13–15 show the left-ear HRTFs in the median, horizontal, and frontal planes for all subjects. Overall a gain of up to 20 dB can be found as well as an attenuation of narrow bands of more than 40 dB. As expected, the HRTFs in general converge to 0 dB for low frequencies. The main differences between directions are found as characteristic peaks and dips above 1 kHz.

For HRTFs in the median plane (Fig. 13) it can be observed that the highest amplitudes—up to 10–15 dB—are seen in the frequency range of 3–5 kHz and for frontal directions. HRTFs for directions behind the subject have in general low amplitudes, especially at the highest frequencies. For most directions it is possible to

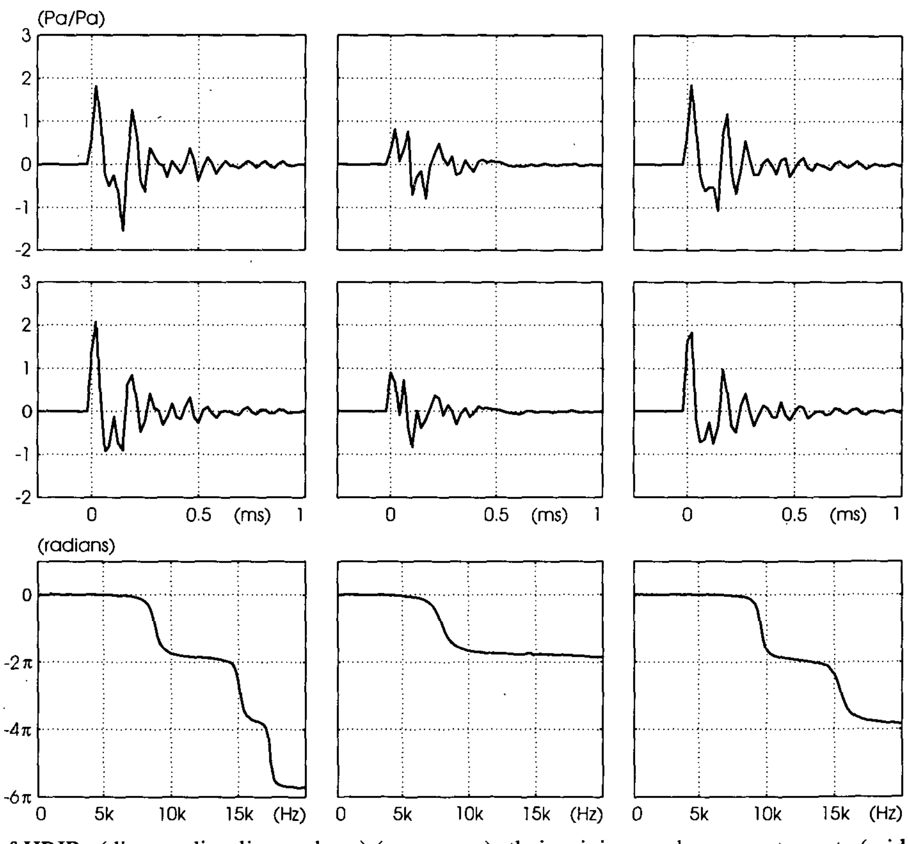

Fig. 12. Examples of HRIRs (disregarding linear phase) (upper row), their minimum-phase counterparts (middle row), and their all-pass phase (lower row). Left column—subject JRH left ear, (ϕ, θ) = (90°, 0°); middle column—subject AKA left ear, (ϕ, θ) = (0°, 0°); right column—subject MAF left ear (ϕ, θ) = (90°, 0°).

find a structure, which is followed by most of the subjects up to 10 kHz. However, a common structure is difficult to find above 5 kHz for the three lowest directions in front, since these HRTFs are characterized by many closely spaced dips.

Differences in HRTFs between angles are large in some zones, while difficult to find in others. An example where the variation with angle is minimal is the upper area, from which the three middle sets of curves in the second column seem quite similar.

Fig. 14 shows the HRTFs for the horizontal plane. The directional variation is considerably larger than for the median plane because of the shadowing effect of the head. This is clearly illustrated by the attenuation of the high frequencies for the ear opposite the sound source. The highest amplitudes, 15–20 dB, are seen around 4 kHz for an azimuth of 45° (upper set of curves in the second column).

The difference between HRTFs of neighboring angles seems to be largest in the shadow zone. As an example, there is a large difference between the HRTFs in the two lower sets of curves in the third column, whereas,

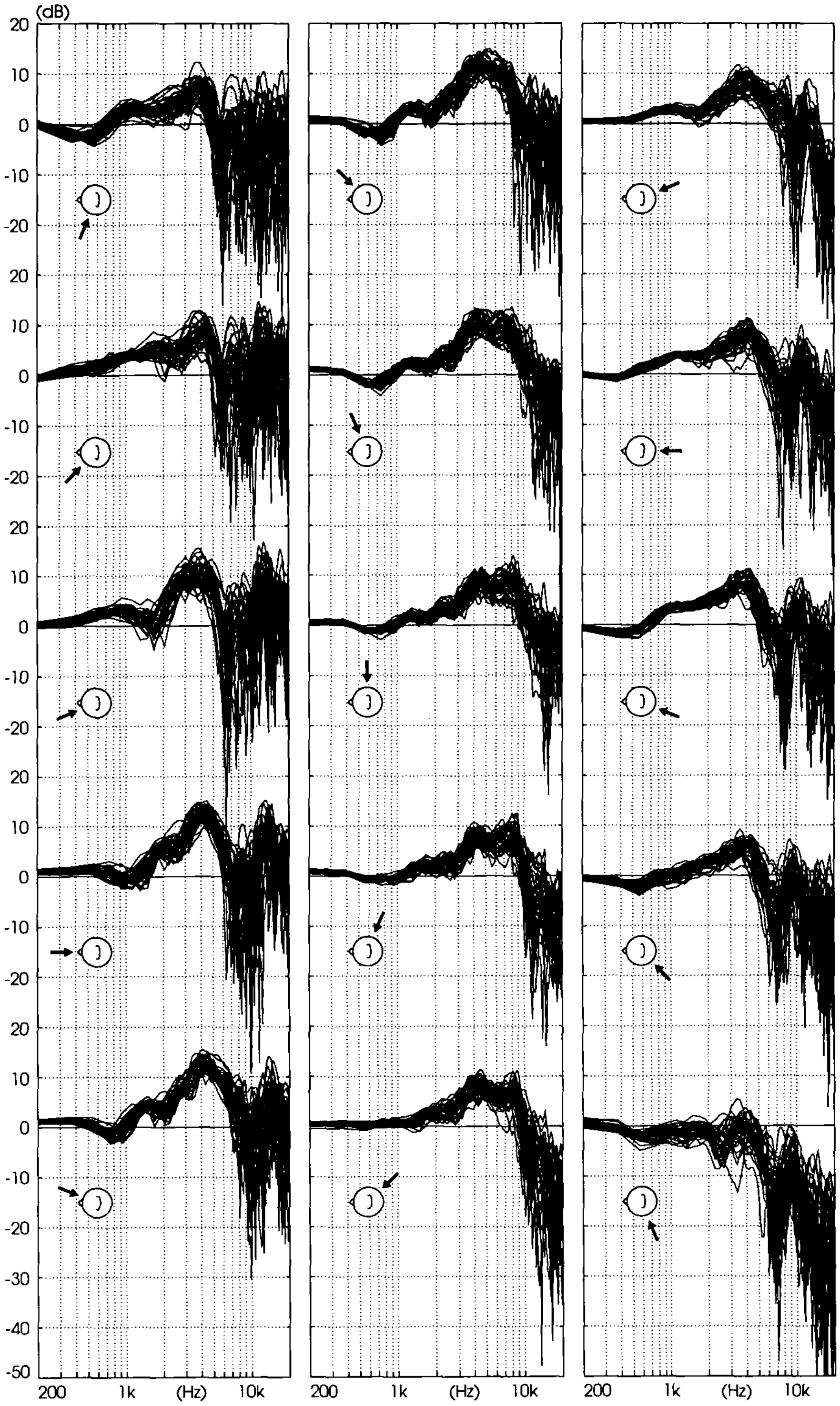

Fig. 13. HRTFs of left ears of all subjects, covering median plane.

for instance, the three middle sets of curves in the second column are more similar.

The HRTFs for the frontal plane are shown in Fig. 15. As already seen in the horizontal plane, there is a general attenuation of the high frequencies for HRTFs on the side opposite the sound source. The HRTFs for directions above the head seem quite similar and with little interindividual variation, whereas the HRTFs for directions below the horizontal plane have more complex and direction-dependent structures. Below the horizontal plane and in the shadow zone, the HRTFs are dominated by highly individual narrow dips extending far down in frequency.

For most directions in the three planes shown, a quite clear common structure for the HRTFs of all subjects is seen up to about 6–8 kHz. Above 8 kHz it is often possible to follow a general structure, but the characteristic peaks and dips seem to be more individual than at lower frequencies.

However, there are certain critical zones where con-

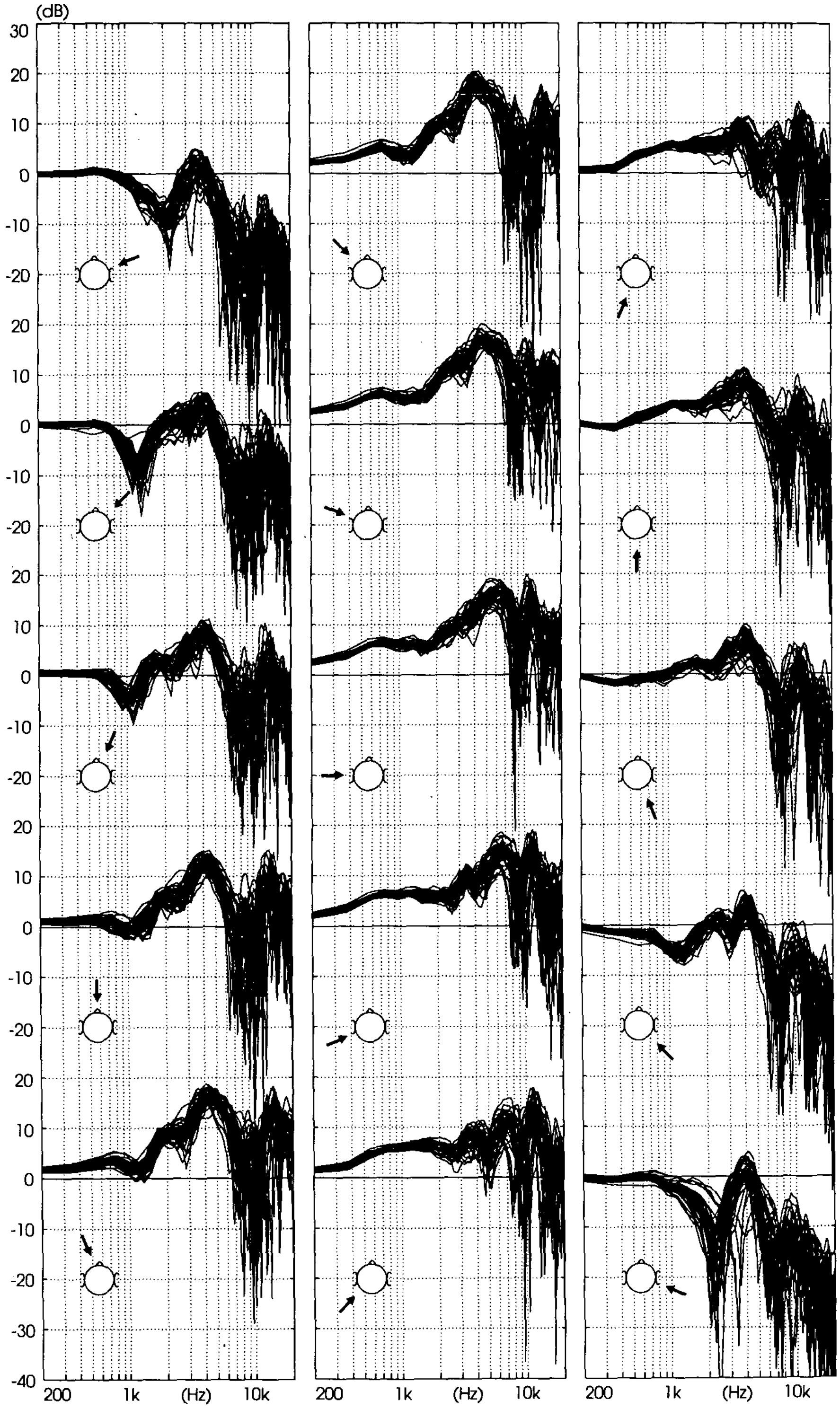

Fig. 14. HRTFs of left ears of all subjects, covering horizontal plane. To economize on space, direction to the right is omitted, since it can be found in Fig. 15 for frontal plane.

siderably larger interindividiaul variations can be observed. The largest interindividual variations are seen for the directions where the sound source is opposite the ear or below the horizontal plane.

2.3 Interaural Time Differences for Three Planes

Fig. 16 shows the ITDs for all subjects in the median, horizontal, and frontal planes. The definition of ITD used is the difference between the arrival times for left and right ears. The ITD is thus negative when the sound arrives at the left ear first. The first sample where an HRIR exceeds 10% of its maximum value defines the arrival time. Note that with the sampling frequency used, the time resolution is approximately 20 μs.

For the median plane, only an asymmetry of the subject's anatomy or an incorrect orientation of the head during the measurement can cause ITDs different from zero. As expected, ITDs are centered around zero. The variation is within ±4 samples, or ±80 μs. Because of the quantization in time it is not possible to get a proper impression of the distribution of the ITDs around zero from Fig. 16. A closer analysis has shown that for most

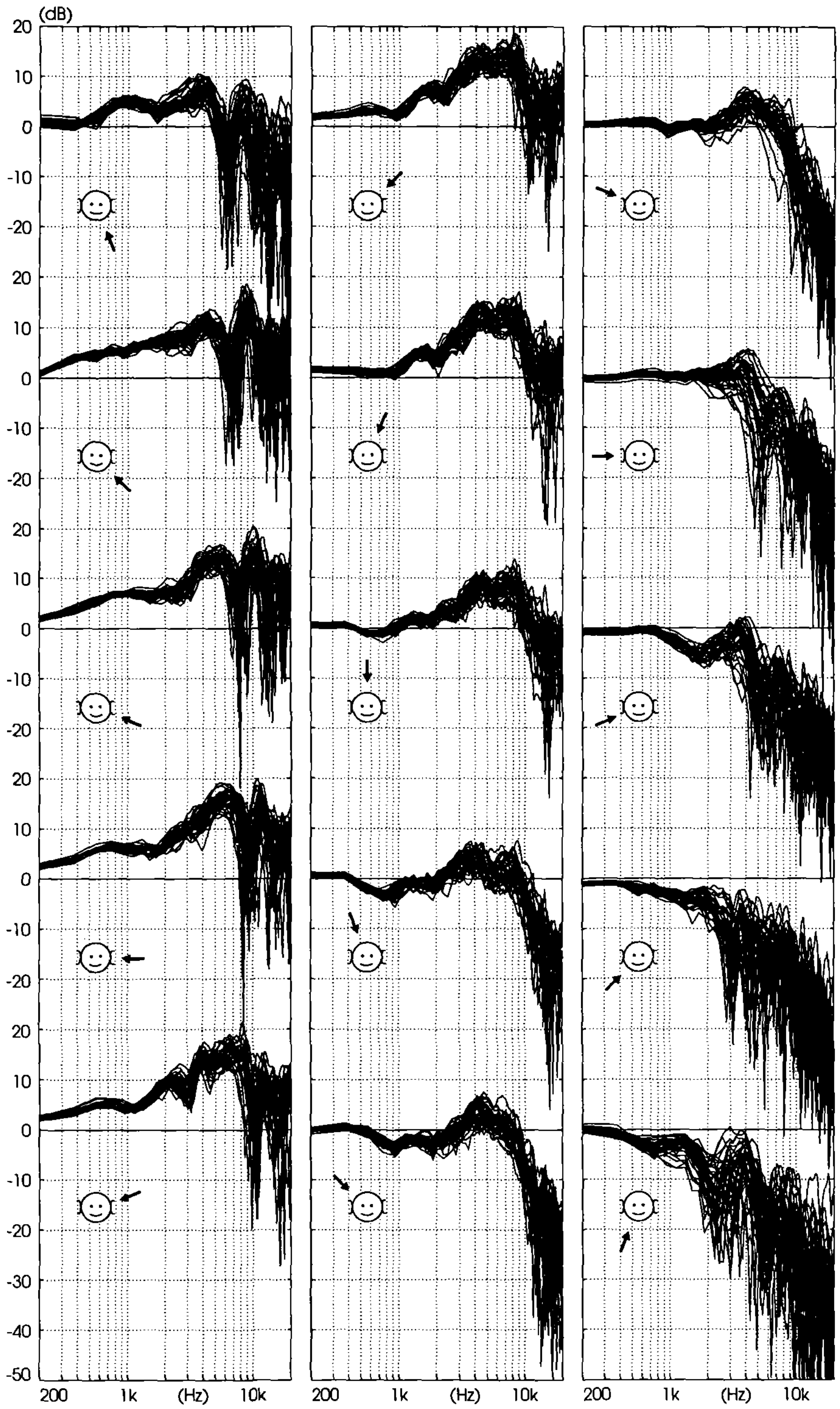

Fig. 15. HRTFs of left ears of all subjects, covering frontal plane.

of the measurements the deviations from zero are below two samples, or 40 μs. These values can be caused by an anatomical asymmetry of 14 mm. In our experience, an asymmetry of this size is not unusual.

The same nonzero values can alternatively be caused by an incorrect orientation of the head by 5° if a head width of 160 mm is assumed. Also an incorrect orientation of this order of magnitude is not unexpected, considering that the subjects' heads were not fixed. Nonzero ITDs for the frontal direction indicate an incorrect azimuth, whereas nonzero values for the direction above the head indicate sideward tilting of the head. Experience from repeated measurements indicates that azimuth errors due to incorrect placement of the position marker prior to the measurements clearly exceed errors due to inaccurate adjustment of the head during the actual measurements.

Errors in elevation due to forward–backward tilting of the head cannot be estimated in a similar procedure. They might exceed the errors mentioned since it was given higher priority to leave the subject's head free of any fixation devices than to control the exact elevation angle.

In the horizontal plane ITDs for the front and back directions are centered around zero, as already seen from the median plane, whereas the maximum and minimum ITD values are seen where one ear faces the sound source.

From the curve for the frontal plane it is seen that ITDs are approximately zero with sound coming from above. The values converge to zero, too, as the direction

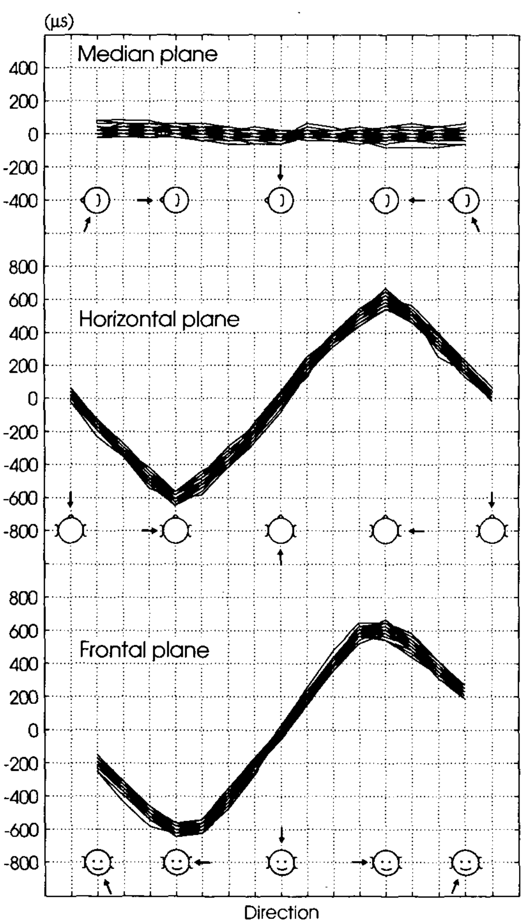

Fig. 16. Interaural time differences for all subjects.

of sound incidence approaches the direction directly below the subject, a direction though where no measurements were taken. Again the extreme values are seen when one ear faces the source.

The ITD for an azimuth of 90° in the horizontal plane is often reported as an indicator of the head size. The mean value for all subjects is 605 μs (mean of absolute values for azimuths of 90° and −90°). A statistical analysis showed that the mean values for women (582 μs) and men (628 μs) were significantly different (0.1% level in a t test).

As the human head in general is not exactly spherical, the maximum ITD value may be found at an angle slightly different from $\phi = \pm 90°$ and $\theta = 0°$, and for a direction where no measurements were made. Manual inspection and interpolation of all individual curves have revealed that the maximum ITD might be some 5–10 μs higher.

A rough impression of the head size can be obtained if the head is considered a sphere with symmetrically placed ears. Seen in this way the mean head diameter would be 162 mm, or 156 mm for women and 176 mm for men. These figures correspond well with our data for head widths.

2.4 Pressure Divisions P_3/P_2

Fig. 17 shows the pressure divisions P_3/P_2 for the left ears of four subjects. Each set of curves shows the five different directions measured. The major information in this figure is that the pressure division for all four subjects is independent of direction, as assumed. Local fluctuations are seen, though, probably due to imperfect replacing of the probe tip and imperfect orientation of the head.

It is also seen from Fig. 17 that pressure divisions are highly individual. For some subjects large fluctuations are seen already from 2 kHz, whereas one subject has a pressure division that is fairly flat up to 7 kHz. Fig. 18 shows pressure divisions for the left ears of all subjects. (For each subject the mean value from the five directions is shown.) A large interindividual variation is seen already from approximately 2 kHz.

As mentioned, the free-field pressure division data from this investigation are used in a parallel work at our laboratory, dealing with the reproduction side of the binaural technique [30].

2.5 Repetition Measurements

To estimate the reliability of the results, three independent measuring series were carried out with one of the subjects on three different days. It should be emphasized that nothing special was done in order to make these measurements more accurate than the others. A comparison of the three sets of measurements indicates the repeatability of the results, including errors due to incorrect orientation of the head (as mentioned in Section 2.3) and incorrect microphone placement.

Fig. 19 shows three repeated HRTF measurements of a subject's left ear for the horizontal plane. Concerning the level at low frequencies the variation is within ±0.5

dB. For the side where the ear faces the sound source, the agreement between the measurements is excellent even up to 20 kHz. The variations between the measurements for the side opposite the sound source are clearly larger. This could be expected since the side opposite the sound source is the most sensitive to changes in direction. Therefore the largest variation due to inaccurate orientation of the subject's head is seen here. However, the variation is still small, except for the extremely sensitive directions in the two lower sets of curves in the right column ($\phi = -112.5°$ and $-135°$).

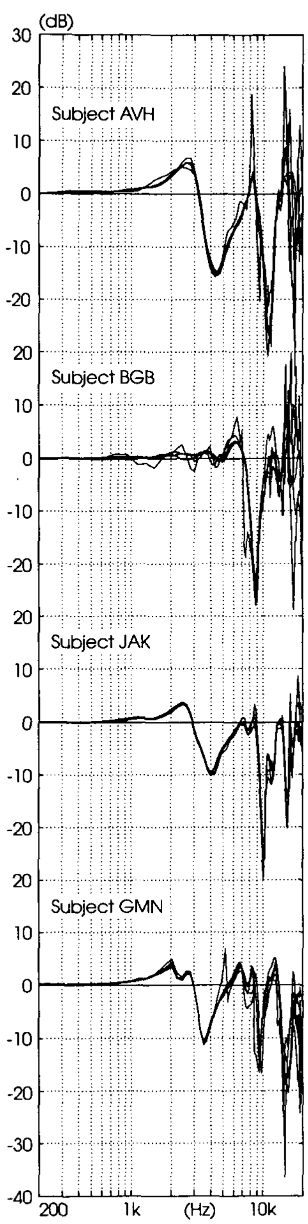

Fig. 17. Examples of pressure division P_3/P_2 for left ears of four subjects. Curves are shown for the five directions (ϕ = 0°, 90°, 180°, and $-90°$ in horizontal plane; θ = 90°).

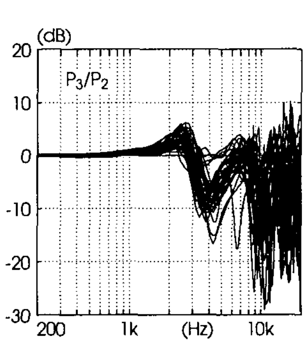

Fig. 18. Pressure divisions for left ears of all 40 subjects.

Probe microphone measurements were also included in the repetition measurements. Fig. 20 shows three HRTF measurements of one subject's left ear, with the probe microphone (left) and the miniature microphone (right). Two things can be seen from Fig. 20. At first, it is obvious that there is a good agreement between the repetition measurements, also for the probe microphone technique. As could be expected, the variation is largest for HRTFs from the side opposite the sound source ($\phi = -90°$).

Second, Fig. 20 can also be used to evaluate the two microphone techniques. In general the two techniques have resulted in almost similar curves. Some of the differences are presumably due to the disturbance of the sound field caused by the probe microphone housing. The largest difference is seen in the shadow side, where the sensitivity to azimuth is high. (Compare the transition from the first frame of Fig. 14 through the frame of the second row, third column of Fig. 15 to the last frame of Fig. 14.) However, in general the differences are small, indicating successful microphone techniques.

Altogether it can be concluded that a comparison of three measurement series indicates that the results are repeatable.

3 DISCUSSION

Data from the present investigation are compared to those obtained in previous studies. The few previous investigations using a blocked ear canal are mentioned in Section 3.1. Comparisons with open ear canal HRTFs are made in Section 3.2 (individual HRTFs) and Section 3.3 (mean HRTFs). The frequency resolution necessary to represent HRTFs is discussed in Section 3.4.

3.1 Previous Studies—Blocked Ear Canal

Wiener and Ross [5] presented measurements from three directions in the horizontal plane. The curves for the three directions are very similar and without the resonances seen in our material. Unfortunately we consider their data unreliable. This is unexpected and in contradiction to their other data on eardrum HRTFs and on the transmission along the ear canal, which are in fair agreement with results of our laboratory [29].

Pösselt et al. [18] showed mean data from measurements with 11 subjects. 21 directions in the median plane were included, of which three are identical to directions included in our measurements. A comparison is given in Fig. 21.

The same structures can be seen in both investigations, but the peaks are higher in the data of Pösselt et al. than in ours. Deviations are also seen at low frequencies for the frontal direction. A high-frequency rolloff—almost the same for all directions—is seen in the data of Pösselt et al. The differences may to some extent be explained by the facts that Pösselt et al. measured 5 mm within the ear canal, that their subjects were sitting, and that their high-frequency "tail" was estimated.

Blocked ear canal measurements were also made by Shaw [34], who presented data from 11 elevation angles. However, the purpose of Shaw's investigation was to evaluate the response of the pinna alone, and effects from the head and the torso were avoided by the use of a sound source mounted near the pinna. Consequently the measurements are not true HRTFs, and the data cannot be compared directly with ours. Nevertheless we have been able to locate many structures from Shaw's data in ours as well, especially at high frequencies, and Shaw's data may serve as a valuable tool in finding the physical structure (torso, head, pinna) that is responsible for the various resonances in the HRTFs.

3.2 Previous Studies—Open Ear Canal, Individual HRTFs

Although many investigations were reported in Section 0.2, only nine of these have presented data in a

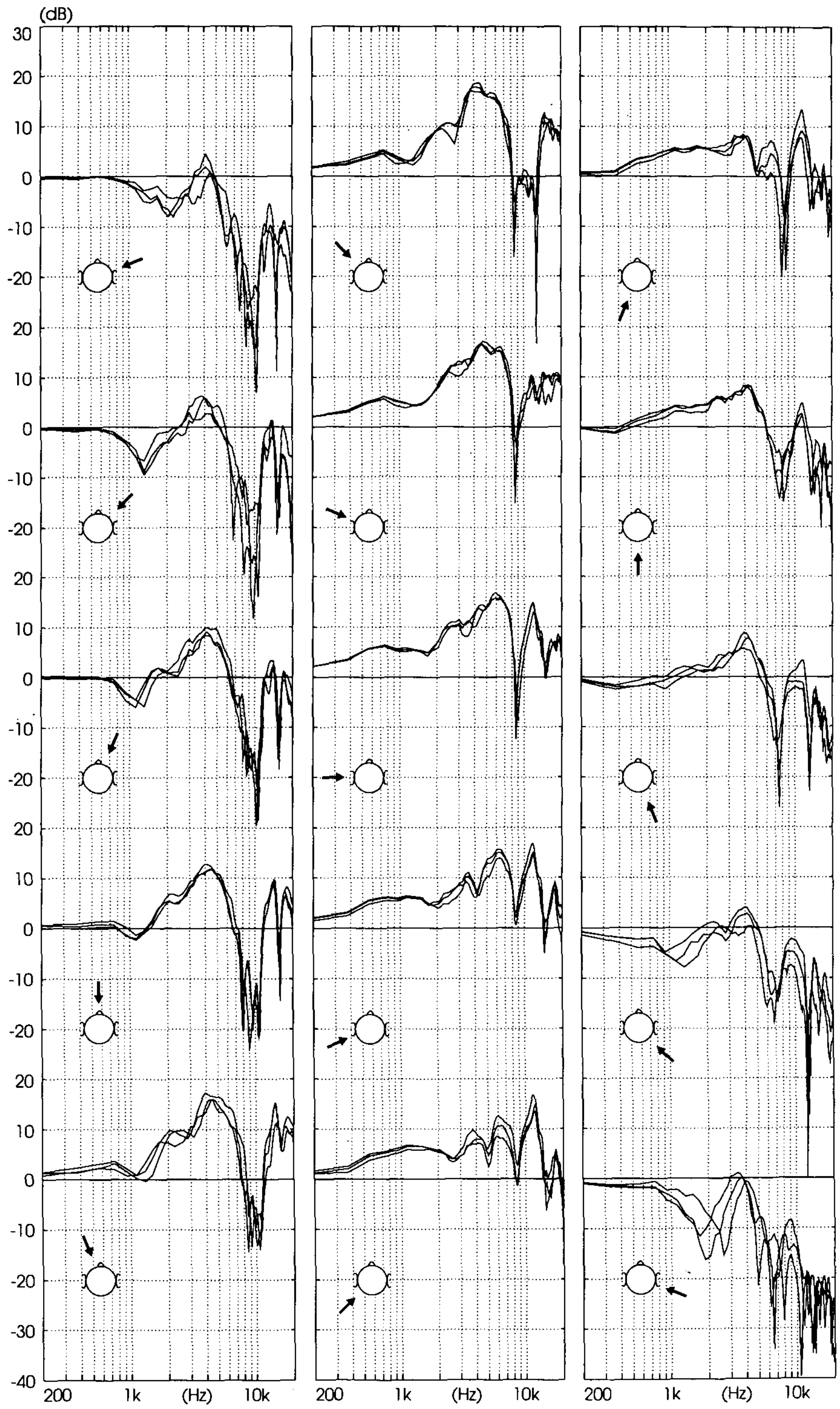

Fig. 19. Three repetition measurements of HRTFs for left ear of subject DOL, shown for horizontal plane.

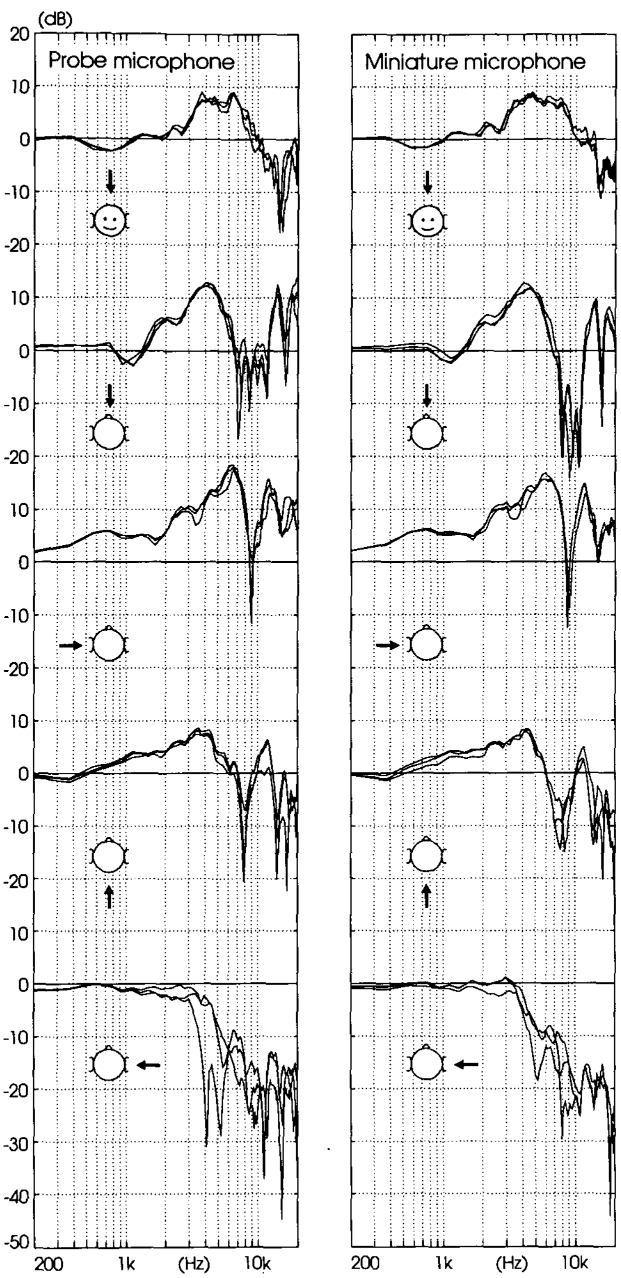

Fig. 20. Three repetition measurements of HRTFs for left ear of subject DOL shown for the five directions measured with both microphone techniques (θ = 0°, 90°, 180°, and −90° in horizontal plane; θ = 90°).

form similar to ours, that is, as individual HRTFs for a group of subjects. Most of the data presented are from directions in the horizontal plane. Wiener and Ross [5] showed data for three subjects, 0° azimuth; Shaw [10] showed left-ear data for 10 subjects, six directions (azimuths 0°, 45°, 90°, 180°, 270°, and 315°); Burkhard and Sachs [12] showed left-ear data for 24 subjects, 90° azimuth; Platte [15] showed left-ear data for four subjects, 90° azimuth; Morimoto and Ando [16] showed left- and right-ear data for three subjects, 13 angles of elevation in the median plane; Genuit [17] showed left-ear data for six subjects, 90° azimuth; Wenzel et al. [19] showed data for both ears of eight subjects, unknown direction; Schmitz and Vorländer [23] showed data for 10 subjects, 0° azimuth; and Hellstrom and Axelsson [27] showed right-ear data for 19 subjects, two directions (azimuths 210°, 270°).

The data of Wiener and Ross [5], Wenzel et al. [19], and Hellstrom and Axelsson [27] were obtained at or close to the eardrum, and cannot be compared with our data since the individual transmission along the ear canal is not known. The data of Morimoto and Ando [16] are presented in a very small figure that cannot be transferred for comparison. Furthermore, though given for the median plane, their data show unexpectedly large differences between left and right ears.

The data of Shaw [10] and Burkhard and Sachs [12] were measured at the entrance to the open ear canal, the data of Platte [15] and Genuit [17] 4 mm inside the ear canal, and the data of Schmitz and Vorländer [23] 4–5 mm inside the ear canal. The reference points in all five investigations are close to the ear canal entrance, and a comparison is reasonable. For the comparisons, open ear canal HRTFs for the present investigation have been obtained by multiplying blocked ear canal HRTFs with the pressure division (the latter calculated for each subject as the mean of five directions). Results at three azimuths are seen in Fig. 22.

In general Shaw's data are in good agreement with ours. For the front direction (left column in Fig. 22) this is true for the level at low frequencies as well as the dip just above 1 kHz. The peak between 2 and 3 kHz is also present in Shaw's data, but it is slightly higher than found in our measurements. At higher frequencies the structures are more difficult to find in Shaw's data, and

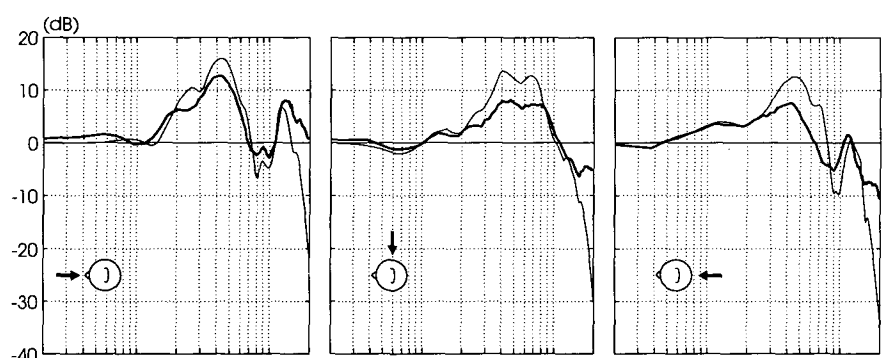

Fig. 21. Comparison of mean blocked ear canal HRTFs of Pösselt et al. [18] (11 subjects, thin lines) and of present investigation (40 subjects, heavy lines).

it is obvious that the dip around 10 kHz is nearly absent. This is probably due to a poor signal-to-noise ratio, a problem Shaw was aware of himself.

For sound from the left (middle column), the agreement is excellent between Shaw's data and those of the present investigation. However, again it is obvious that the signal-to-noise ratio has limited the depth of the 9-kHz dip in Shaw's data.

For the back direction (right column) peaks around 2.5 kHz and at 12–14 kHz, and dips at 3–4 kHz and 8–10 kHz are seen in both investigations. Also the structures at low frequencies are in fair agreement. The frequency range of 4–8 kHz is difficult to compare due to the large interindividual variation, but the levels do agree.

Schmitz and Vorländer presented only data for the frontal direction. The general shape of their data is in fair agreement with that of our data. However, at low frequencies fluctuations are seen in Schmitz and Vor-

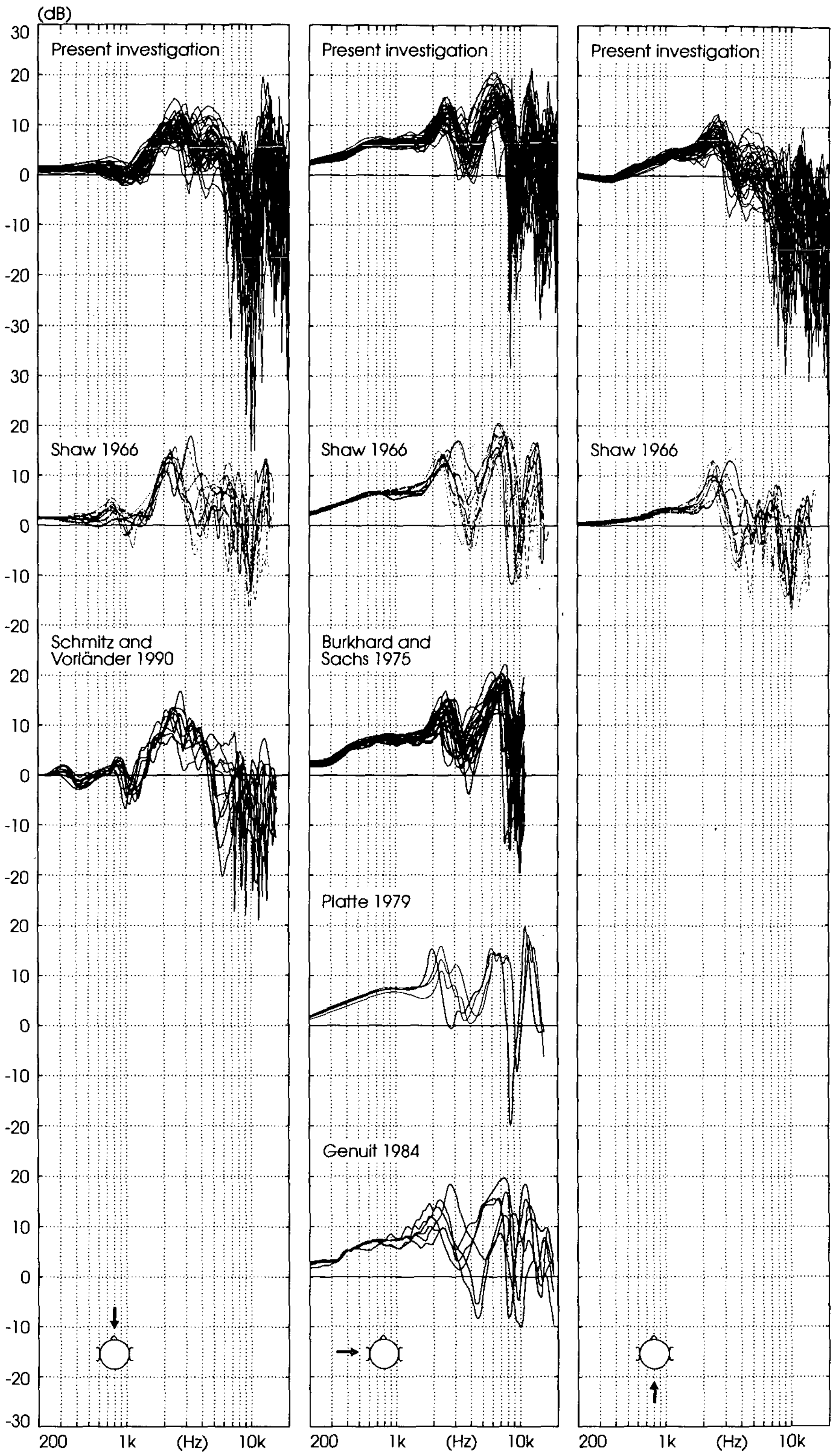

Fig. 22. Comparison of individual open ear canal HRTFs of present investigation, Shaw [10], Burkhard and Sachs [12], Platte [15], Genuit [17], and Schmitz and Vorländer [23].

länder's data, whereas our curves are almost flat up to about 1 kHz. The dip slightly above 1 kHz is somewhat deeper than seen from our data. The differences may be due to a reflection from the subjects' knees, since the subjects were seated in the investigation by Schmitz and Vorländer. The peak between 2 and 3 kHz is present with approximately the same level in both sets of curves, but the slightly lower peak around 5 kHz seen for most of our subjects is not found in Schmitz and Vorländer's data. At high frequencies the dip at 8–12 kHz and the peak at 11–13 kHz in our data are replaced by highly individual dips in Schmitz and Vorländer's data.

Data obtained by Burkhard and Sachs for sound from the left were originally presented for men and women separately. However, for comparison, data for all their 24 subjects are shown together. An excellent agreement is seen with our data. Peaks around 2.5 and 7 kHz, and dips around 4 and 9 kHz, as well as the same low-frequency structure are seen in both investigations.

Platte's data, also for the left direction, show a good agreement with our data. Only the peak around 7 kHz has a different shape for all his subjects, almost as if the top had been cut off. A similar phenomenon is not seen in any of the other investigations.

A fair agreement is also seen with Genuit's measurements, although for two of his persons the 3–4-kHz dip seems to have moved to 4.5 kHz and become much deeper.

Despite the small differences it can be concluded that there is good agreement between our data and the data on individual subjects found in the literature.

3.3 Previous Studies—Open Ear Canal, Mean HRTFs

HRTFs have also been presented in the form of mean values for a number of subjects. A very comprehensive study of this kind was carried out by Shaw [35], who combined HRTF measurements from six investigations [5]–[10] and iteratively adjusted the results to achieve consistency with data on monaural HRTFs [36]–[39] and data on interaural transfer functions [40]. Results are given as HRTFs with a reference point at the eardrum, but in order to obtain this reference point, some data [6], [10] had to be transferred from the entrance of the ear canal to the eardrum. For this transmission a curve was fitted to data from two investigations [5], [28]. Other data [8] were transformed from outside the ear canal by a modified version of this curve. When the numbers of subjects and angles in the various investigations contributing to Shaw's curves are taken into account, it is seen that a majority of the data have been transferred in this manner along the ear canal. We have therefore taken the liberty of transferring Shaw's mean data (back) to the ear canal entrance, using his data for the transmission along the ear canal [35, fig. 2].

In the upper row of Fig. 23 Shaw's data (thin lines, taken from a later tabulated version [41] of the original curves) are compared with our mean values (heavy lines). The following rows show comparisons with later studies not included in Shaw's data (and using almost the same measurement point): Blauert [2], [3] (12 subjects); Mehrgardt and Mellert [13] (20 subjects); Platte [15] (6 subjects); Genuit [17] (6 subjects); and Okabe and Miura [24] (28 subjects). The method used to obtain the mean differs slightly among investigations, a fact that is discussed later in this section.

An excellent agreement is seen between Shaw's data and ours. This is especially true with sound incidence from left and from behind. For these directions only very small differences are seen in the location of peaks and dips. For the front direction the differences are somewhat larger, and above 6 kHz the level found by Shaw is 3–4 dB higher than what we found, a fact that caused the 10-kHz dip to be less deep. Blauert's mean data and ours are close, except for minor deviations at low frequencies and around the notch at 10 kHz.

A similar agreement is not seen between the mean values presented by Mehrgardt and Mellert and ours. For the front direction Mehrgardt and Mellert's data have a clear dip slightly above 1 kHz. In our data only a shallow dip around the same frequency is seen. In the range of 2–6 kHz differences of up to 8 dB are seen, whereas the dip at 10 kHz is in perfect agreement. With left sound incidence Mehrgardt and Mellert obtained approximately the same structures as we did. The peak at about 2.5 kHz, though, is some 4 dB higher than in our data, and the peak we found at about 6.5 kHz is placed near 8 kHz in their data. The dip at 10 kHz is about 3 dB deeper than the one located at 9 kHz in our data. With sound incidence from the back, the peak at about 2.5 kHz is found to be some 3 dB higher in Mehrgardt and Mellert's data. Above 4 kHz the structure is quite different in the two investigations, except that both curves have a dip near 10 kHz.

Platte's results are dominated by local fluctuations, which are not seen in the other investigations. Disregarding these, a fair agreement is seen with our data. The same general structures are seen, and most of the peaks and dips are located at approximately the same frequencies in both sets of data. At some of the resonances, though, the levels differ significantly.

In general Genuit's mean data are close to ours, although the two high-frequency notches are less deep. This may be partly due to a lower signal-to-noise ratio and partly due to a larger spread between subjects of the exact notch frequency.

A good agreement is seen between the data of Okabe and Miura and our data. Differences are seen up to 3 kHz, but for higher frequencies the two sets of data are almost identical. It is worth noting that these two investigations are the single studies with the largest number of subjects.

Some of the differences between investigations can be explained by the different techniques that were used for averaging between subjects. Our mean values were calculated on a sound-level (decibel) basis, frequency by frequency. The same applies to the mean values of Blauert and of Genuit. The method used by Shaw is similar in nature, but somewhat more complicated due to the different form of the original data he used.

Characteristic for the "frequency by frequency" averaging is that narrow dips and peaks located at slightly different frequencies for each subject will be flattened and appear more shallow. Consequently the average will not represent a typical subject. Mehrgardt and Mellert have tried to overcome this problem by shifting each individual HRTF along the logarithmic frequency axis to give the best fit before averaging. The idea that the structures would fit better between individuals, following this frequency shift, corresponds to the assumption that humans are equally proportioned but with different sizes.

This is a very rough simplification, and it might be better to identify each structure of the HRTFs by parameters such as frequency, Q, level, etc. and to average these parameters separately (parametric averaging). Platte made an approach to this by modeling individual HRTFs with an electronic circuit and averaging the components.

Some impact of the averaging method is seen in the data presented. For the back direction the investigations by Mehrgardt and Mellert and by Platte showed a "dip–peak" structure at 5–8 kHz which is not seen in Shaw's and our mean data. Although the level is somewhat uncertain, a similar structure can be identified for some *individuals* in Shaw's and our data (Fig. 22, right column). Also a few dips have become slightly deeper and possibly more representative of a typical subject.

In general, though, neither Mehrgardt and Mellert's data nor Platte's data seem more similar to the individual data seen in Fig. 22. Mehrgardt and Mellert did not publish their individual data, and it is not possible to see whether the discrepancies are also seen at this level. Platte's average data can be compared with his own individual data (as included in our Fig. 22), but even on this background his average does not seem to be especially representative of a typical subject.

The lack of improvement does not disqualify parametric averaging, but it emphasizes the need for clearly identifiable structures present for all subjects before averaging. Such an agreement between subjects seems to be lacking, especially in the 2–7-kHz frequency range.

A way of reducing the variation between subjects is to choose another measurement point. As mentioned in Section 0.3, the reason for choosing the blocked ear canal for our measurements was the low variation between subjects at this point, as shown by Hammershøi and Møller [29].

The left column of Fig. 24 gives examples of individual and mean HRTFs for an open ear canal from the present investigation. Especially in the frequency range

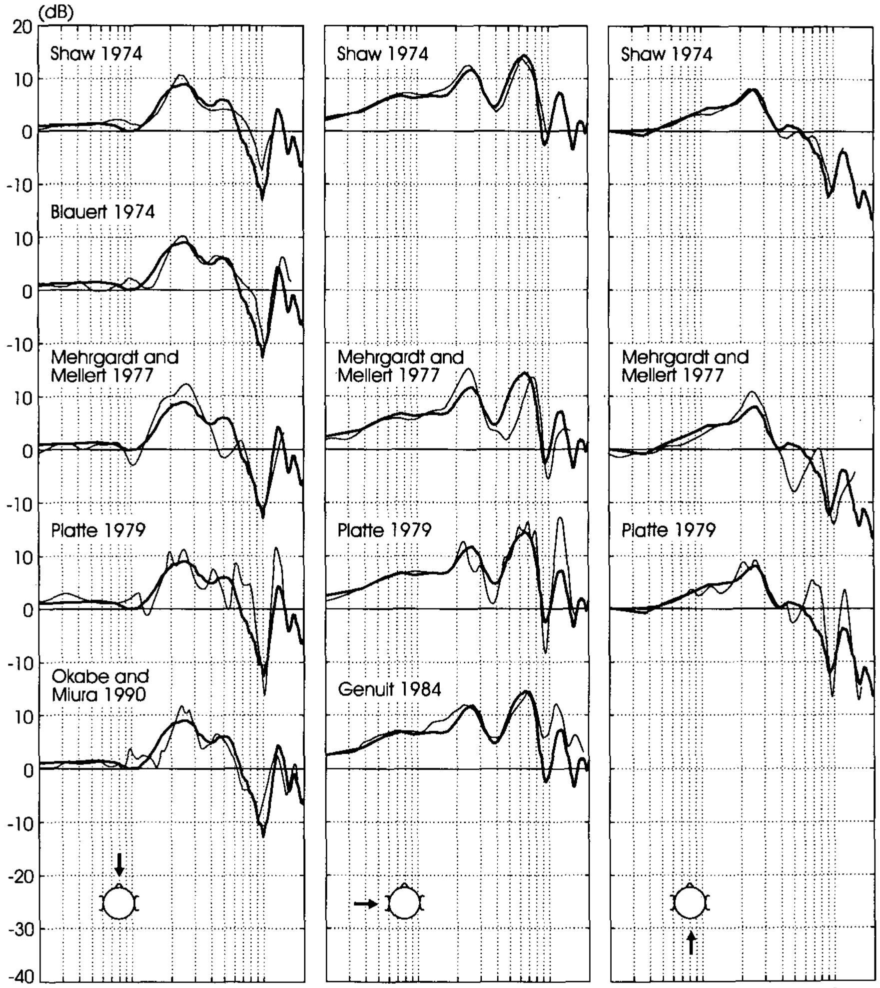

Fig. 23. Comparison of mean open ear canal HRTFs of present investigation (heavy lines) with data of Shaw [35], [41] (literature review); Blauert [2], [3] (12 subjects); Mehrgardt and Mellert [13] (20 subjects); Platte [15] (6 subjects); Genuit [17] (6 subjects), and Okabe and Miura [24] (28 subjects).

of 2–7 kHz, where there is high intersubject variation, mean curves are misleading. The right column of the figure shows blocked ear canal data. The variation between subjects is clearly reduced, and the mean value gives a better representation of a typical subject. An exception is the frequency range of 7–12 kHz for the frontal direction, where the variation between individuals is still very high, even for the blocked ear canal data.

3.4 Frequency Resolution

The measurement technique used in the present investigation provides knowledge about the fine structures of the HRTFs, such as narrow dips and peaks. Some previous investigations have presented HRTFs as one-third-octave values only. Fig. 25 illustrates clearly the lack of information with such a coarse frequency resolution. It is obvious that information is completely lost about the narrow dips around 9 and 14 kHz and about the peak around 12 kHz.

4 CONCLUSION

The measurement technique with miniature microphones mounted in earplugs is simple and has proven useful for measuring head-related transfer functions. Video monitoring the subject's head has given good accuracy in the angle of sound incidence without disturbance from supporting devices. Use of the MLS technique has made it possible to present results in the frequency domain as well as in the time domain. The synchronous sampling in the two channels has secured reliability, also when interaural time differences are concerned. Even when the arrival time has been accounted for, the HRTFs do not in general represent minimum-phase systems.

The choice of reference point at the blocked ear canal entrance has proven rational. Earlier investigations, showing that transmission from the reference point and further into the ear canal is one-dimensional, have been supported. The variation between subjects of HRTFs measured at the blocked ear canal is smaller than at the same point with open ear canal. The directional characteristics of the sound transmission to the ear are thus best described by blocked ear canal HRTFs.

Direction-dependent peaks and dips are seen in the HRTFs, and it is encouraging to the binaural technique in general that an interindividual variation in these characteristics for most directions is small up to approximately 8 kHz. Above 8 kHz it is still possible to find a general structure for most directions, but the variation has increased considerably. In general the data indicate an objective basis for the binaural technique.

For comparison with most results of earlier investigations it has been necessary to adjust our data to open ear canal HRTFs. Despite the larger variation in open ear canal data, good agreement was found with other investigations.

The representation of HRTFs as an average across subjects is somewhat problematic, especially for open ear canal HRTFs. Narrow peaks and dips occurring at slightly different frequencies for each subject are flattened. The curves become more shallow and thus do not represent a typical subject. More sophisticated averaging techniques from the literature have not at present given significant improvements at this point.

As a natural continuation of the present and the parallel work [30], both concerning objective aspects of the binaural technique, a large series of psychoacoustic experiments have been carried out at our laboratory [42], [43]. These experiments were intended to answer funda-

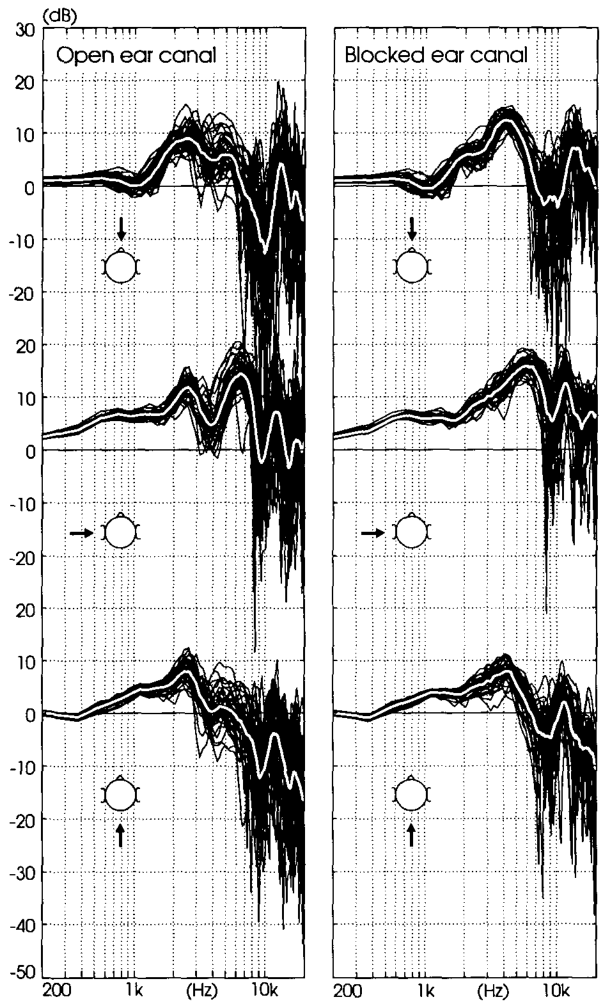

Fig. 24. Comparison of left ear HRTFs for open ear canal (left column) and blocked ear canal (right column). White curves represent means.

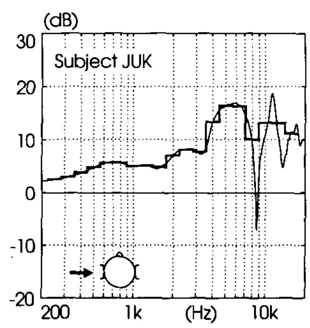

Fig. 25. Example of an open ear canal HRTF shown with one-third-octave frequency resolution and with frequency resolution of present study.

mental questions regarding subjective evaluation of the binaural technique. Data from the present investigation have also been used to compute design goals for headphones intended for playback of traditional recordings [44].

5 ACKNOWLEDGMENT

Economic support from Brüel & Kjær A/S, Perceptive Acoustics A/S, the National Agency of Industry and Trade, Denmark, Aalborg University, and the Danish Technical Research Council is greatly acknowledged. The authors would like to thank their former colleagues Kim Alan Larsen and Jørn Vagn Hundebøll for their participation when the experiments were planned and carried out. Kim Alan Larsen was a coauthor, too, when the investigation was preliminarily reported [1]. We would also like to thank Anne Kirstine Andersen, our coordinator, for her handling of all appointments with the test subjects. Also Claus Vestergaard, the technical employee at our laboratory, deserves appreciation for his valuable help in the practical work. At last we would like to thank all our patient subjects for standing hour after hour in the anechoic chamber with various microphones in their ears.

6 REFERENCES

[1] D. Hammershøi, H. Møller, M. F. Sørensen, and K. A. Larsen, "Head-Related Transfer Functions: Measurements on 40 Human Subjects," presented at the 92nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 40, p. 443 (1992 May), preprint 3289.

[2] J. Blauert, *Räumliches Hören* (S. Hirzel Verlag, Stuttgart, Germany, 1974).

[3] J. Blauert, *Spatial Hearing—The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA, 1983).

[4] H. Møller, "Fundamentals of Binaural Technology," *Appl. Acoust.*, vol. 36, pp. 171–218 (1992).

[5] F. M. Wiener and D. A. Ross, "The Pressure Distribution in the Auditory Canal in a Progressive Sound Field," *J. Acoust. Soc. Am.*, vol. 18, pp. 401–408 (1946).

[6] F. M. Wiener, "On the Diffraction of a Progessive Sound Wave by the Human Head," *J. Acoust. Soc. Am.*, vol. 19, pp. 143–146 (1947).

[7] Z. Yamaguchi and N. Sushi, "Real Ear Response of Receivers," *J. Acoust. Soc. Jpn.*, vol. 12, pp. 8–13 (1956).

[8] D. W. Robinson and L. S. Whittle, "The Loudness of Directional Sound Fields," *Acustica*, vol. 10, pp. 74–80 (1960).

[9] G. Jahn, "Über den Unterschied zwischen den Kurven gleicher Lautstärke in der ebenen Welle und im diffusen Schallfeld," *Hochfrequenztech. Elektroakust.*, vol. 69, pp. 75–81 (1960).

[10] E. A. G. Shaw, "Ear Canal Pressure Generated by a Free Sound Field," *J. Acoust. Soc. Am.*, vol. 39, pp. 465–470 (1966).

[11] C. L. Searle, L. D. Braida, D. R. Cuddy, and M. F. Davis, "Binaural Pinna Disparity: Another Localization Cue," *J. Acoust. Soc. Am.*, vol. 57, pp. 448–455 (1975).

[12] M. D. Burkhard and R. M. Sachs, "Anthropometric Manikin for Acoustic Research," *J. Acoust. Soc. Am.*, vol. 58, pp. 214–220 (1975).

[13] S. Mehrgardt and V. Mellert, "Transformation Characteristics of the External Human Ear," *J. Acoust. Soc. Am.*, vol. 61, pp. 1567–1576 (1977).

[14] H.-J. Platte and P. Laws, "Technische Probleme beim Einsatz kopfbezogener stereofoner Übertragungsverfahren," *Rundfunktech. Mitt.*, vol. 22, pp. 22–26 (1978).

[15] H.-J. Platte, "Zur Bedeutung der Aussenohrübertragungseigenschaften für den Nachrichtempfänger 'menschliches Gehör,'" Ph.D. dissertation, Rheinisch-Westfälischen Technischen Hochschule, Aachen, Germany (1979).

[16] M. Morimoto and Y. Ando, "On the Simulation of Sound Localization," *J. Acoust. Soc. Jpn.* (E), vol. 1, pp. 167–174 (1980).

[17] K. Genuit, "Ein Modell zur Beschreibung von Aussenohrübertragungseigenschaften," Ph.D. dissertation, Rheinisch-Westfälischen Technischen Hochschule, Aachen, Germany (1984).

[18] C. Pösselt, J. Schröter, M. Opitz, P. L. Divenyi, and J. Blauert, "Generation of Binaural Signals for Research and Home Entertainment," in *Proc. 12th Int. Cong. on Acoustics* (Toronto, Ont., Canada, 1986), vol. 1, B 1-6.

[19] E. M. Wenzel, F. L. Wightman, and S. H. Foster, "A Virtual Display System for Conveying Three-Dimensional Acoustic Information," in *Proc. Human Factors Society, 32nd Ann. Meet.* (1988), pp. 86–90.

[20] F. L. Wightman and D. J. Kistler, "Headphone Simulation of Free-Field Listening, I: Stimulus Synthesis," *J. Acoust. Soc. Am.*, vol. 85, pp. 858–867 (1989).

[21] J. C. Middlebrooks, J. C. Makous, and D. M. Green, "Directional Sensitivity of Sound-Pressure Levels in the Human Ear Canal," *J. Acoust. Soc. Am.*, vol. 86, pp. 89–108 (1989).

[22] J. C. Middlebrooks and D. M. Green, "Directional Dependence of Interaural Envelope Delays," *J. Acoust. Soc. Am.*, vol. 87, pp. 2149–2162 (1990).

[23] A. Schmitz and M. Vorländer, "Messung von Aussenohrstossantworten mit Maximalfolgen-Hadamard-Transformation und deren Anwendung bei Inversionsversuchen," *Acustica*, vol. 71, pp. 257–268 (1990).

[24] K. Okabe and T. Miura, "Construction of a Head and Torso Simulator for Orthostereophonic System (OSS)" in Japanese, *J. Acoust. Soc. Jpn.*, vol. 46, pp. 885–892 (1990).

[25] T. Sone, Y. Suzuki, S. Takana, and K. Suzuki, "Distance Perception in Sound Localization and Its Control by Simulation of Head-Related Transfer Functions," presented at the 14th International Congress on Acoustics, Beijing (1992 Sept.), paper L7-4.

[26] J. C. Middlebrooks, "Narrow-Band Sound Localization Related to External Ear Acoustics," *J. Acoust. Soc. Am.*, vol. 92, pp. 2607–2624 (1992).

[27] P.-A. Hellstrom and A. Axelsson, "Miniature Microphone Probe Tube Measurements in the External Auditory Canal," *J. Acoust. Soc. Am.*, vol. 93, pp. 907–919 (1993).

[28] G. Djupesland and J. J. Zwislocki, "Sound Pressure Distribution in the Outer Ear," *Scand. Audiol.*, vol. 1, pp. 197–203 (1972).

[29] D. Hammershøi and H. Møller, "Sound Transmission to and within the Human Ear Canal," submitted to *J. Acoust. Soc. Am.*

[30] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen, "Transfer Characteristics of Headphones Measured on Human Ears," *J. Audio Eng. Soc.*, vol. 43, pp. 203–207 (1995 Apr.).

[31] D. D. Rife and J. Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences," *J. Audio Eng. Soc.*, vol. 37, pp. 419–444 (1989 June).

[32] J. Sandvad and D. Hammershøi, "Binaural Auralization: Comparison of FIR and IIR Filter Representation of HIRs," presented at the 96th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 42, p. 395 (1994 May), preprint 3862.

[33] D. Hammershøi and J. Sandvad, "Binaural Auralization: Simulating Free-Field Conditions by Headphones," presented at the 96th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 42, p. 395 (1994 May), preprint 3863.

[34] E. A. G. Shaw, "The External Ear: New Knowledge," *Scand. Audiol.*, Suppl. 5, pp. 24–50 (1975).

[35] E. A. G. Shaw, "Transformation of Sound Pressure Level from the Free Field to the Eardrum in the Horizontal Plane," *J. Acoust. Soc. Am.*, vol. 56, pp. 1848–1861 (1974).

[36] L. J. Sivian and S. D. White, "On Minimum Audible Fields," *J. Acoust. Soc. Am.*, vol. 4, pp. 288–331 (1933).

[37] G. Jahn and S. Vogelsang, "Die einohrige Richtcharakteristik des Ohres," *Hochfrequenztech. Elektroakust.*, vol. 68, pp. 50–56 (1959).

[38] W. Schirmer, "Die Richtcharakteristik des Ohres," *Hochfrequenztech. Elektroakust.*, vol. 72, pp. 39–48 (1966).

[39] J. Blauert, "Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf," Ph.D. dissertation, Rheinisch-Westfälischen Technischen Hochschule, Aachen, Germany (1969).

[40] W. E. Feddersen, T. T. Sandel, D. C. Teas, and L. A. Jeffress, "Localization of High Frequency Tones," *J. Acoust. Soc. Am.*, vol. 29, pp. 988–991 (1957).

[41] E. A. G. Shaw and M. M. Vaillancourt, "Transformation of Sound-Pressure Level from the Free Field to the Eardrum Presented in Numerical Form," *J. Acoust. Soc. Am.*, vol. 78, pp. 1120–1123 (1985).

[42] H. Møller, C. B. Jensen, M. F. Sørensen, and D. Hammershøi, "How Well Do Existing Binaural Systems Work?", in preparation.

[43] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural Technique: Do We Need Individual Recordings?", in preparation.

[44] H. Møller, C. B. Jensen, D. Hammershøi, and M. F. Sørensen, "Design Criteria for Headphones," *J. Audio Eng. Soc.*, vol. 43, pp. 208–232 (1995 Apr.).

Biographies of the authors were published in the 1995 April issue of the *Journal*.