Evaluation of artificial heads in listening tests

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

Published in:
Journal of the Audio Engineering Society

Publication date:
1999

Evaluation of Artificial Heads in Listening Tests*

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

Acoustics Laboratory, Aalborg University, DK-9220 Aalborg Ø, Denmark

The localization performance was studied when subjects listened 1) to a real sound field and 2) to artificial-head recordings of the same sound field. The experiments took place in a standard listening room where each stimulus (female speech) was emitted from one of 19 loudspeakers, and the subjects were to indicate the perceived sound source. The artificial-head recordings were made a) by the artificial heads' built-in microphones and b) by blocked ear canal microphones. The reproduction was carried out by carefully equalized headphones. Eight artificial heads were included in the investigation, and 20 subjects participated, except for the experiment with recordings from built-in microphones, which was performed for eight subjects. When compared to real life, the localization performance with the artificial heads resulted in an increased number of errors independent of the recording technique. In general, the directions in the median plane were frequently confused, not only with nearby directions, but also with directions further away. For some artificial heads there was also an increase in confusions of directions outside the median plane. A much better performance is obtainable with binaural recordings made in the ears of humans. This encourages the design and production of improved artificial heads.

0 INTRODUCTION

During the last decade the binaural technique has been used increasingly for recordings of noise and music. The main reason is the capability of the technique to store and reproduce three-dimensional aspects of a sound field by means of only two channels.

The idea behind the binaural technique is the following. The input to the hearing consists of two signals—the sound pressures at the eardrums. If these signals are recorded in the ears of a listener and reproduced exactly as they were (usually through headphones), then the complete auditory impression is recreated, including spatial aspects such as direction and distance to sound sources.

In most practical applications recordings cannot be made in the listener's own ears, and during recording the listener is usually replaced by a copy of a human head, called artificial head or dummy head. Several artificial heads are commercially available, and they do produce recordings with a surprisingly natural reproduction. Nevertheless, it is often claimed that the three-dimensional space is distorted, and that localization errors occur. Often mentioned imperfections are: confusions between front and back, elevation of sound sources, sound sources perceived too close, and in-the-head localization. It seems that these shortcomings have limited the success of the artificial-head recording technique.

0.1 Previous Investigations

The proficiency of artificial heads in giving a true reproduction of the auditory space has been assessed in several listening experiments. In the investigation by Damaske and Wagener [1] three localization experiments were made. In the first experiment seven subjects indicated the apparent source position out of 11 possible equidistant source positions in the upper median plane, while lying on one side on a table in the anechoic chamber. The signal was speech produced by a human speaker during the test, and the subject was instructed to keep his eyes closed at all times and to indicate the apparent source directions by calling out integer numbers corresponding to loudspeaker positions. Twenty-five responses were collected for each direction, seemingly in total for the seven subjects. In the second experiment two trained

* Presented at the 102nd Convention of the Audio Engineering Society, Munich, Germany, 1997 March 22–25; revised 1999 January 15.

subjects listened to the reproduction of signals picked up by probe microphones in the ears of an exhibition manikin head, which had replaced the listener's head in the anechoic chamber. The signals were reproduced (instantly) by Telefunken T 50 headphones, for which no equalization strategy is reported. The localization performance was significantly deteriorated with the artificial-head signals. In the third experiment recordings for the full horizontal plane (still with 18° spacing) were used. Two trained subjects participated, and a total of 18 responses were collected for each direction. Localization errors and inversions were frequent (39% inversions for azimuth 216°, and 33% inversions for azimuth 144°). (Comparable data were obtained in an investigation by Damaske and Mellert [2], where the same artificial head and setup were used, but the recordings were reproduced by loudspeakers using cross-talk cancellation).

In the investigation by Wilkens [3] localization tests were made using sources in the horizontal plane (15° spacing including 0° azimuth). The test signal was a 100-ms noise burst, highpass filtered at 2 kHz. The signal was produced by loudspeakers in a darkened anechoic chamber. Both a real-life experiment and an experiment with recordings from an artificial head (seemingly a precursor of Neumann KU80) were made. No report is given about the type of headphone or equalization. The subjects knew the physical setup and could only report on directions corresponding to the loudspeaker positions. Two listeners participated and the stimuli were repeated 16 times per direction. In real life, the absolute angle was overestimated for azimuths between 15° and 75° and between 90° and 120°. With artificial-head recordings the same happened from 15° to 75°, whereas the angle was slightly underestimated for directions from 120° to 150°. Inversions were observed, although they never exceeded 10% for any direction.

Laws and Platte [4] made localization tests with 12 source directions in the horizontal plane (30° spacing including 0° azimuth). (Later reports by Platte and Laws [5] and by Platte [6] describe the same experiments, and some details have been taken from these.) Test signals were speech emitted by equalized loudspeakers in an anechoic room. The localization performance was studied in real life and when listening to recordings made with a Neumann KU80 artificial head. The headphone used for reproduction seems to be a Beyer DT 302, but no information is given about equalization. Eight subjects participated, and the stimuli were repeated five times in random order. The subjects were sitting blindfolded in the dark setup during both conditions and they gave their responses in terms of integer numbers of the clock, corresponding to the positions of the loudspeakers. In real life, the mean responded direction was close to the source, but rather large deviations (up to 60° for the mean) were seen with the artificial-head recordings. For the sources at 0° and 180° there were no confusions between front and back in real life but 52% front-to-back and 12.5% back-to-front confusions with the recordings. For sources at 0° and 180° subjects were also given the opportunity to report an "in-the-head" perception, and the number of these increased from 5% in real life to the 11% with the recordings.

In the investigation by Boerger et al. [7] localization tests were carried out using recordings made in a recording studio with an artificial head, presumably a Neumann KU80. Test signals were speech emitted from loudspeakers at 6 directions in the horizontal plane (azimuths of −20°, 0°, 20°, 160°, 180°, −160°). Details on headphone and equalization are not reported. Five subjects participated, and fractions of front-to-back and back-to-front inversions were approximately 28% and 47%, respectively. Details, for example, the response method and the distribution of inversions on directions, are not reported, and no real-life experiments were made.

In the investigation reported by Poulsen [8] localization tests are reported, using recordings of 12 equidistant source directions in the horizontal plane (30° spacing including 0° azimuth). Experiments with recordings from two artificial heads (Neumann KU80 and Knowles Kemar) as well as an experiment with the real-life sources were included. Three listeners participated, and the stimuli—noise and speech—were each repeated four times for each direction and presented in random order. The recordings were reproduced by headphones (the Peerless PMB 6, the AKG 140, and the TDH 39) for which no equalization strategy was applied. No localization errors were observed in real life, whereas many confusions occurred with the reproduction of the artificial-head recordings. On the average 28% and 41% inversions and 33% and 29% neighbor confusions are reported for the KU80 and the Kemar, respectively.

Genuit and Platte [9] carried out localization tests using 12 source directions in the horizontal plane. Experiments were made in real life and with recordings from two artificial heads, a Neumann KU80 and an IENT 77 (presumably the construction of the author's laboratory as described by Laws and Platte [10]). Stimuli were wideband sound emitted by loudspeakers in an anechoic room (random order). No information is given on headphone and equalization, nor on number of subjects and repetitions. Only few errors were seen in real life. For both heads, around 50% errors were seen for the sources at 0° and 180°. For the remaining directions, a high number of errors was seen for KU80, somewhat fewer for the IENT 77.

Genuit [11] made localization tests using 12 source directions in the horizontal plane. Real-life experiments as well as experiments with recordings from three artificial heads, a Neumann KU80 and two of the author's laboratory, denoted IENT 81 and IENT 84, the latter with a simplified geometry. (The IENT 81 and IENT 84 are presumably the two heads later manufactured by HEAD acoustics GmbH, and denoted HMS I and HMS II, respectively, see Section 0.3.3.) Stimuli were sinusoidally modulated (10 Hz) white noise emitted from loudspeakers in an anechoic room. No information is given on headphone and equalization, nor on number of subjects and repetitions. Best results, that is, almost no errors, were obtained in real life. For the KU80, a large number of errors, in particular inversions, were seen for the 0° and 180° directions (56%) as well as outside the median plane. For the IENT 81 results were close to those of real life, while slightly more errors were seen for the IENT 84.

In the investigation by Theile and Spikofski [12] localization tests were made using recordings of 12 equidistant source directions in the horizontal plane (30° spacing including 0° azimuth). The stimulus was speech and the subjects gave their responses in terms of integer numbers of the clock, 26 or more listeners participated. Only experiments with recordings from artificial heads (Neumann KU81 and an IENT head, presumably the head denoted IENT 81 in [11] and presumably later manufactered as HMS I) were included. The KU81 had a diffuse-field equalization and the IENT head had a free-field equalization, and the equalization of the headphones—STAX lambda pro—was switched to suit the equalization of the dummy head used. On the average both heads resulted in approximately 45% confusions. The fraction of front-to-back inversions was on the average 16% for the IENT head and 20% for the KU81. For the particular source at 0° the fraction of front-to-back confusions was approximately 5% for the IENT head, and 25% for the KU81, whereas the fraction of back-to-front confusions for 180° was approximately 50% for both heads.

Genuit [13] made localization tests using seven source directions in the upper median plane, equally spaced by 30° and including 0° elevation. Experiments were made in real life and with recordings from a detailed artificial head (presumably the head denoted IENT 81 in [11] and presumably later manufactured as HMS I). Stimuli were synthetic wide-band bell tones emitted from loudspeakers in an anechoic room. Seven trained subjects participated. Nothing is reported on headphone and equalization, nor on experimental design. The number of directional errors increased from 5.3% in real life to 12.9% with the artificial-head recordings. The fraction of "percept close to head" increased correspondingly from 4.6% to 10.7%.

Other investigations have assessed the validity of using artificial heads for sound recording and reproduction by comparative listening tests (e.g., Krumbacher [14] for other qualities of the spatial perception and syllable identification, Plenge [15] for the phenomenon of in-head localization, Wettshurek [16] for absolute difference limen in directional hearing in the median plane, Plenge [17] for discussions on lateralization and localization, Toole [18] for evaluation of loudspeaker systems). However, the objectives and methods for these investigations differ much from those just mentioned and those of the present study, thus they are not referenced in detail here.

In conclusion, the study of previous investigations confirms that reproduction of artificial-head recordings leads to an impaired localization compared to real life. Due to the differences in experimental conditions, way of specifying and presenting results, computing statistics (only few have employed statistical comparisons at all), and—in particular—the very limited number of subjects in most of the previous investigations, we decline from further comparative conclusions at this point.

0.2 Aim of Investigation

It is the purpose of this investigation to study the quality of artificial-head recordings. Note that the investigation thus constitutes an evaluation of the binaural technique when the recordings are made with artificial heads. We have previously carried out similar evaluations of the technique with recordings made in human ears [19], [20].

0.3 Artificial Heads Included in Investigation

Seven different artificial heads from five producers were included in the investigation. Some of them could be modified in various details, which made a total of eleven different configurations.

0.3.1 Knowles Electronics, Inc.

The Knowles Electronics Mannikin for Acoustic Research (KEMAR) consists of a head and a torso originally designed from anthropometrical data and for the purpose of general acoustic research. It includes ear simulators according to IEC 711 [21] (and ANSI S3.25 [22]) for which the sound pressure at the microphone (Brüel & Kjær 4134) is intended to correspond approximately to the sound pressure existing at the human eardrum. Four different humanlike pinnae are provided, ordered from smaller to larger (two sides): DB060/061, DB065/066, DB090/091, DB095/096. In the following the KEMAR equipped with these pinnae (in the same order) are denoted by KEMAR 1, KEMAR 2, KEMAR 3, and KEMAR 4.

0.3.2 Georg Neumann GmbH

Both of Georg Neumann GmbH's artificial heads, KU80i and KU81i, have heads and pinnae with a humanlike shape. As the only ones in the investigation they do not include a torso. The heads have a 4-mm ear canal leading to acoustical networks and microphones (KM83 for KU80i and KK83 for KU81i). The KU80i is designed to simulate the recording of the sound at a certain point in a human ear canal, whereas the acoustical network of KU81i is designed for equalizing the frequency response to a flat response for diffuse sound incidence.

It should be noted that at the time of the investigation Neumann had announced their artificial head KU100. Unfortunately the company was not able to lend us a sample for a few days to make recordings. The head was not yet commercially available, hence we were not able to include it in the investigation.

0.3.3 HEAD Acoustics GmbH

The two artificial heads of Head Acoustics GmbH are marketed as parts of complete recording and playback systems. The heads are very different. They both include head, shoulder, and pinnae, but the head of the recording system HMS I is a replica of a specific human, whereas the head of the recording system HMS II is a model of very simplified geometry. Even the pinnae have been highly simplified. For both heads, a box with electronic equipment makes up the part of the torso below the shoulders. These heads have a 4-mm ear canal terminated by the recording microphones Brüel & Kjær 4166 (HMS I) and 4165 (HMS II). The HMS I head was constructed at the Rheinisch-Westfälische Technische Hochschule, Aachen, Germany, and it is sometimes called the "Aachener Kopf."

The playback systems (HPS II for HMS I, HPS III for HMS II) include playback equalizers to be used with a Stax SR lambda professional headphone. Recording and playback equalizers have settings for free-field equalization as well as a so-called "independent-of-direction (ID)" equalization.

0.3.4 Brüel & Kjær

Brüel & Kjær's two artificial heads have the same outer shape, consisting of a head and a torso, both of simplified geometry, and equipped with humanlike pinnae. The 4128 includes ear simulators according to IEC 711 [21] (and ANSI S3.25 [22]) in which the simulated eardrum pressure is recorded by a Brüel & Kjær 4134 microphone. At the 5930 no attempt is made to simulate the ear canal, and studio microphones 4009 are mounted nearly flush with the ear canal entrance.

0.3.5 University of Toronto

This head is designed at the Institute of Biomedical Engineering at the University of Toronto. It has a humanlike shape, includes humanlike pinnae (DB065/066 from a KEMAR), and fits onto a KEMAR torso, with which it was used in the present investigation. It was originally made for the measurement of earplug attenuation and hence has an accurate simulation of the ear canal tissue, but no microphones are provided with the head. For the purpose of this investigation miniature microphones (Sennheiser KE 4-211-2) were mounted in the ear canals, flush with the ear canal entrances. The head is only produced on request and has no type number; hence in the following it is simply denoted by Toronto.

1 METHOD

By means of psychoacoustic experiments the localization performance was evaluated in real life and when listening to artificial-head recordings of the same sound field. The experimental method was in most respects identical to that used in our previous investigation with binaural recordings made in the ears of human subjects [19]. The description will therefore be confined to a summary of the procedures, supplemented by complete descriptions of issues specific to the present investigation. More details may thus be found in our earlier report.

The experiments included three listening conditions: 1) listening in real life, 2) listening to recordings made with the built-in microphones of the artificial heads, and 3) listening to recordings made with miniature microphones mounted flush with the entrance to the blocked ear canal of the artificial heads.

The latter condition constitutes a situation where the recordings are influenced only by the "outer" geometry of the head. The fact that this results in recordings with full spatial information was shown theoretically by Møller [23] and verified from measurements by Hammershøi and Møller [24] and in listening tests by Møller et al. [19].

1.1 Setup and Procedure

The experiments were carried out in a standard listening room [25], where 19 loudspeakers were located around the subject. Of these, 14 were positioned in various directions on a sphere with a radius of 1 m, with seven of them being in the median plane. The remaining five were in more distant positions. The subjects listened to a 5-s recording of a female voice, either directly from the loudspeakers or indirectly as an artificial-head recording, made in the same setup and reproduced by means of headphones. The subject was sitting in the setup in both situations and kept his or her head still during the stimuli. The loudspeakers were visible to the subjects, and the experiments were carried out as identification experiments, where the subjects responded from which loudspeaker they perceived the sound.

Due to practical circumstances a few more curtains were present in the listening room during the recordings than during the experiments. Hence the reverberation time was slightly lower (approximately 0.3 s in contrast to the normal approximately 0.4 s). The possible consequence of this difference is discussed in Section 2.3, and it is verified that it did not affect the results. Hence the difference is ignored elsewhere in this paper.

1.2 Subjects

Twenty paid students with controlled normal hearing participated as listeners, ten of each sex, aged 20–30 years. They were all skilled in psychoacoustic experiments, but they were not in any way selected for their hearing or localization proficiency. The same subjects participated in a previous investigation on the selection of a typical human subject for binaural recordings [20]. Prior to that eight of the subjects had participated in our investigation on individual and nonindividual binaural recordings [19].

Some of the experiments (experiments B, Ax, and Ay) were only carried out for eight of the subjects. These subjects were the same eight subjects who had participated in the first experiment [19]. The real-life performance of these eight subjects did not deviate from the real-life performance of the other twelve subjects [20, app.].

1.3 Recordings

Two sets of recordings were made for each artificial head: recordings with the built-in microphones, and recordings with miniature microphones mounted flush with the entrance to the blocked ear canal. The type of miniature microphones and the mounting technique were the same as used earlier for measurements of head-related transfer functions on human subjects [26] and for making binaural recordings in the ears of human subjects [19]. Since the outer geometry is identical for the 4128 and 5930 heads, only one set of miniature microphone recordings was made for these. Recordings with the built-in microphones of the KEMAR were only made with one size of pinnae—the KEMAR 2.

The procedure used during recording was identical to that used in the earlier experiment with recordings in human ears [19], except that a Panasonic SV-3700 DAT recorder was used for intermediate storage before the recordings were put onto the disk. The recordings with built-in microphones of the HMS I head were made including the recording equalizer of the system in position "free-field," whereas the recordings with the HMS II head were made without the recording equalizer of the system.

1.4 Reproduction and Equalization

Most of the manufacturers of artificial heads gave (and still give) only very unspecific instructions about the reproduction, which was then left to the experimenters' decision.

Headphones were chosen because of their almost complete channel separation. The total transmission should then include equalization for 1) the transfer function of the recording microphones and 2) the transfer function of the headphone measured at the point in the ear canal where the recording is made [23]. The equalization was accomplished by digital IIR filters implemented at the signal processor board during reproduction.

1.4.1 Recordings with Built-in Microphones

There is in general some uncertainty related to which point in the human ear canal that the sound recorded by the built-in microphone represents. It was therefore considered necessary to make the measurement of the headphone transfer function at the artificial head. It was further decided to use an FEC headphone (a headphone with free-air equivalent coupling to the ambient; see Møller et al. [27]). It was shown earlier that this procedure leads to a cancellation of effects from a possibly incorrect simulation of the ear canal and eardrum [19, app.].

A Beyerdynamics DT990 Professional headphone was chosen. We showed earlier that this headphone has approximate FEC properties [27]. The headphone was fitted onto each artificial head, and the combined transfer function of the recording microphone and the headphone was found by measuring the transfer function from headphone input to microphone output. The inverse of the amplitude response was used as target in a Yule–Walker IIR filter design (maximum 32nd order, carried out in MATLAB, The MathWorks Inc.). The measured transfer functions were in general not minimum-phase transfer functions, and the equalization was thus only correct with regard to the amplitude response, since the inverted all-pass sections were not accounted for.

A few modifications of the target responses were necessary to stabilize and optimize the digital filters. Lack of dc transmission in the headphone transfer function would result in an infinite dc gain of the target response. Dc values were therefore manually inserted in the target responses to obtain fairly flat responses at low frequencies. A corresponding problem exists close to the half sampling frequency, where the target responses were manually flattened. Manual modifications were also needed where the amplitude response of the headphone had a narrow dip, resulting in a very high and sharp peak in the target response.

The KU80i had a severe lack of high-frequency output, and in order to avoid an extremely high filter gain at high frequencies, the target response was limited by a third-order Butterworth low-pass filter at 6.5 kHz. This resulted in an effective restriction of the frequency range for this head.

As mentioned earlier, the manufacturer specified complete playback methods for the HMS I and HMS II artificial heads. A Stax SR lambda professional headphone was to be used in combination with playback filters of the reproduction platforms HPS II and HPS III, respectively. These directions on headphone and equalizing filters were also used for the experiments, but for practical reasons the filters were implemented at the signal processor board in the same way as the filters for the other recording heads. Since the recordings for the HMS II head were made without the recording equalizer, the implemented filters include this equalizer as well. The magnitude of the HPS II playback equalizer in the position "free-field" was obtained from the manufacturer in graphic form, whereas the data for the HMS II recording and the HPS III playback equalizers (also position "free-field") were obtained from the manufacturer in terms of magnitude at 1/24-octave frequencies for the frequency range of 20 Hz to 20 kHz.

For completeness it should be mentioned that the manufacturer specifies recordings with KU81i to be reproduced without extra equalization using a diffuse-field equalized headphone. We showed earlier that the frequency responses of commercial headphones with a claimed specific equalization vary a lot [27], and none of the headphones included in that investigation fulfilled the design goal to even a reasonable extent. The general equalization procedure described here was therefore used also for KU81i.

Fig. 1 shows the target responses and the implemented IIR filters for all artificial heads for the left and right ears. The figure also shows the target responses which would have been used for the HMS I and HMS II heads and the Stax headphone, had the general equalization procedure been used for these heads as well.

1.4.2 Recordings at Blocked Ear Canal

Here the recording point was well defined and the equalization could be based on headphone transfer functions measured at the individual subjects. Also here it was decided to use an FEC headphone. It had been shown earlier [19, app.] that this procedure minimizes the error which is inevitably introduced by the use of nonindividual recordings, such as artificial-head recordings.

The Beyerdynamics DT990 Professional headphone was used also for these recordings. The design of the equalization filters was carried out as described in the preceding section, except that the headphone transfer functions were measured on the human subjects with the miniature microphones mounted at the blocked ear canal, and filters were designed individually for each subject. The filters were reported earlier ([19, fig. 4] eight subjects, [20] twelve subjects).

1.5 Experimental Design

As mentioned already there were three experiments, one of real-life listening and two with listening to artificial-head recordings accomplished with two different microphone techniques. In addition to the microphone technique, the latter two experiments differed in their statistical focus: one was carried out for a small group of subjects and with several repetitions in order to favor the power of statistical tests, whereas the other was carried out without repetitions but for a larger group of subjects to improve the general validity of the results.

1.5.1 Experiment A: Real Life

Each subject listened to each loudspeaker six times. The experiment was divided into two sessions with three repetitions each. The stimulus order was random in each session. The sessions had a duration of approximately 10 min, and they were separated by a short break. The number of stimuli for each subject was 114, giving a total of 2280 for the 20 subjects.

(The real-life experiment was the very same experiment that was used and reported earlier for the whole group of subjects [20], and for the group of eight used in the present experiments B, Ax, and Ay [19].)

1.5.2 Experiment B: Recordings with Built-in Microphones

Each subject listened to each loudspeaker six times for each of the eight artificial heads included in this experiment. For a particular subject the experiment was carried out on four separate days. The experiments on a particular day were divided into two parts by a long break. Each part covered recordings from one head.

For each head the experiment was divided into two sessions with three repetitions in each session. The stimulus order was random in each session. The sessions had a duration of approximately 10 min, and they were separated by a short break. Only eight subjects participated in this experiment. The order of heads was randomized among the subjects. The number of stimuli per subject was 912, giving a total of 7296 for all subjects.

1.5.3 Experiment C: Recordings at Blocked Ear Canal

Each subject listened to each loudspeaker only once for each of the 10 artificial heads included in this experiment. For a particular subject the experiment was divided into three sessions, separated by short breaks. Each session had a duration of approximately 10 min. For each subject the stimulus order was random for the whole experiment. The number of stimuli per subject was 190, giving a total of 3800 stimuli for all subjects.

1.5.4 Experiments Ax and Ay: Real-Life Control Experiments

The experiments included a considerable number of stimuli for each subject, and even when no feedback was given, the possibility of a learning effect should be considered. To facilitate an assessment of this, an extra real-life experiment, experiment Ax, was carried out for

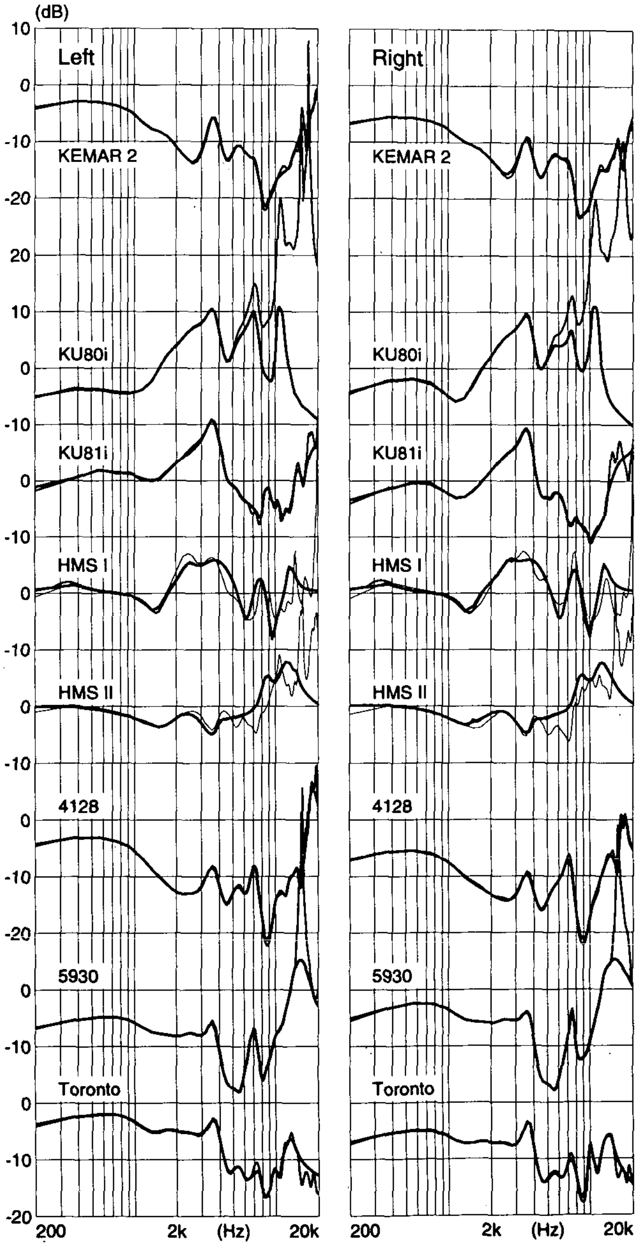

Fig. 1. Equalization filters for reproduction of recordings with built-in microphones. — target; ▬ implemented IIR filter. Headphones were Beyerdynamics DT990 Professional (except for HMS I and HMS II) or Stax SR lambda professional (for HMS I and HMS II). Thin lines indicate targets for HMS I and HMS II with the Stax headphone, had the general equalization procedure been used (see text) (for HMS I taking into account that the recording was made including the HMS I recording equalizer).

eight of the subjects on a separate day at a late stage in the course of the study. At this time each of these subjects had participated in experiment A of the present study, experiments B, C, and D of our earlier study [19], experiment B of our earlier study [20], and experiments B and Ay of the present study (in the order mentioned).

Another possible source of error is the slightly different reverberation time of the listening room during recording and during the experiments. To facilitate an assessment of possible effects of this, an extra real-life experiment, experiment Ay, was carried out with the same reverberation time as during the recordings. This experiment was carried out for eight of the subjects on a separate day just prior to experiment Ax.

Both experiments Ax and Ay were carried out like experiment A. The number of stimuli per subject was 114, giving a total of 912 for the eight subjects (for each of Ax and Ay).

2 RESULTS AND DISCUSSION

The way of presenting the results and performing statistical analysis will to a great extent follow the procedures used in our previous studies. Some details are therefore described only briefly, and more information is found in [19].

For statistical analysis errors are classified into four groups. If a response is given at another cone of confusion than where the stimulus was given, it is called an *out-of-cone error*. A response at the correct cone but at an incorrect direction is called a *within-cone error*, except when stimulus and response are in the median plane, in which case it is designated a *median-plane error*. A response given in the same direction as the stimulus, but at an incorrect distance, is called a *distance error*.

With the present experimental design, the number of errors in a certain category will follow a binomial distribution. The null hypothesis assumes that the error probability is the same for the two conditions under test. The test function follows a hypergeometrical distribution, and the test is called a Fisher–Irwin test (see, for example, [28]). One-sided tests are used whenever the sign of a possible difference can be anticipated. In order to give the most powerful tests, only stimuli that actually can lead to errors in a certain category are included in each test, and in the calculation of error percentages. Also analyses of variance are carried out for those results that fulfill the preconditions for such analyses (error percentages for median-plane and distance errors).

Results from real-life listening are reported in Section 2.1. Section 2.2 gives the results from listening to artificial-head recordings made with built-in microphones, including a comparison with real-life listening. Section 2.3 presents an evaluation of two possible sources of error: learning and deviation of reverberation time during recording. The results with blocked ear canal microphones are given in Section 2.4, including a comparison with real life. The two microphone techniques are compared in Section 2.5, whereas detailed results for each single head are given in Section 2.6. Section 2.7 presents a comparison between the present results with artificial-head recordings and earlier results with nonindividual recordings made in human ears. The analyses of variance are given in Section 2.8. Some further comments are given in Section 2.9.

2.1 Real Life

The results of real-life listening (experiment A) are shown in Fig. 2. The abscissa gives the stimulus position and the responded position is given as the ordinate. Black circles represent answers, and the area of each circle is proportional to the number of answers for the particular combination of stimulus and response.

Correct answers are found on the diagonal, and most of the responses are indeed seen here. However, it is also obvious that the subjects do not localize sound sources perfectly. The major part of the errors are seen for sources in the median plane. Directions in the upper median plane (FRONT HIGH, ABOVE, and BACK HIGH) are frequently confused, and sound coming from FRONT LOW and BACK LOW are sometimes perceived at various other directions in the median plane. Also wrong judgment of distance is a common error.

An important observation is that sound sources in the FRONT direction are almost always perceived in the correct direction. The same applies to the source at BACK.

2.2 Recordings with Built-In Microphones

The results from artificial-head recordings with built-in microphones are seen in Fig. 3. The figure is given for all heads together. Thus the observations need not be valid for every single head.

It is obvious that considerably more errors are made than in real life. Much more confusions are seen between the upper median-plane directions (FRONT HIGH, ABOVE, and BACK HIGH), and a remarkable increase is seen in the number of errors for the low median-plane sources (FRONT LOW, BACK LOW).

Additional directions have also come up with errors, of which some should be mentioned. In real life there were almost no errors for the sound sources in the FRONT direction (except for distance errors), whereas—for artificial-head recordings—these sources are frequently perceived in other directions in the median plane, quite often in the BACK direction. Correspondingly, the sound source in the BACK direction is frequently perceived in other median-plane directions, quite often in the BACK HIGH, ABOVE, or FRONT HIGH directions.

More errors are also seen for all directions outside the median plane, except for LEFT and RIGHT, where no errors at all are seen for real life and for the artificial heads. It is believed that these directions are identified merely by the interaural time difference cue.

Figures similar to Fig. 3 can be presented for each artificial head. However, in order to save space we have chosen to present detailed information for each head only from experiment C, since experiment C included more subjects and thus a wider range of "fits" between humans and artificial heads. Detailed results for single heads are presented in Section 2.6.

At this place the difference between each artificial head and real life is analyzed statistically for each error category, and the results are given in numerical form in Table 1 and in graphical form in Fig. 4.

It is seen that there is a higher percentage of errors for all heads and for all error categories. For median-plane errors the increase is significant at the 0.1% level for all heads. The increase in out-of-cone errors and within-cone errors is significant at various levels for most of the heads. The same applies to distance errors. However, for distance errors an important observation which affects the statistical tests is reported in Section 2.3, so the reader is requested to ignore this for the moment. (For this reason the significance indications are put in parentheses.)

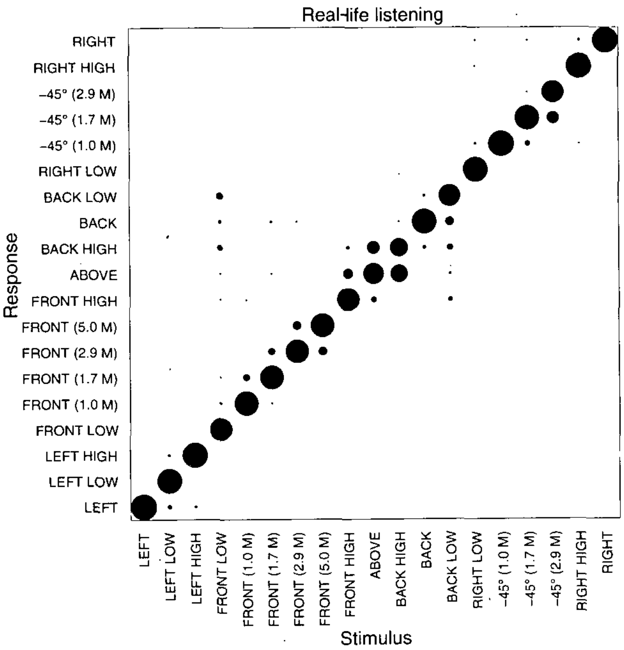

Fig. 2. Real-life listening (experiment A, 2280 stimuli). The area of each circle is proportional to the number of answers for the particular combination of stimulus and response. A full circle [e.g., at (LEFT, LEFT)] corresponds to the total number of stimuli for each direction (which is 120 in this case).

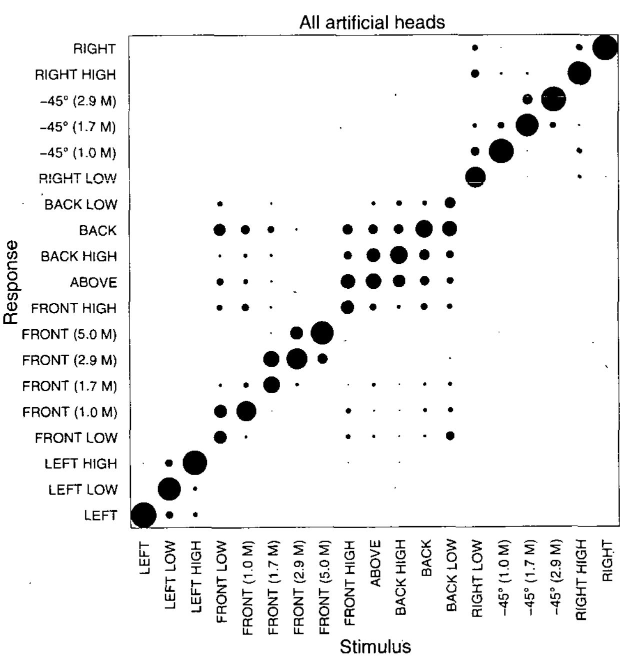

Fig. 3. Artificial-head recordings (experiment B, all heads, 7296 stimuli).

2.3 Effect of Learning and Reverberation Time

A possible effect of learning was studied in a repeated real-life experiment carried out at a late stage in the course of the study. Table 2 shows a comparison of real-life performance in the initial experiment (experiment A) and in the late experiment (experiment Ax).

A striking decrease is observed for median-plane errors (significant at the 0.1% level). This indicates a learning effect for these errors. An equally prominent increase in distance errors is seen (significant at the 0.1% level). The latter is quite surprising. As an explanation one might suspect that the decrease of median-plane errors (and also out-of-cone and within-cone errors) results in more potential distance errors. A detailed analysis has shown that this is not the reason, since the reduction of median-plane errors (plus out-of-cone and within-cone errors) hardly affects the directions where distance errors are possible (FRONT and −45°). We have no other explanation for the changes of errors with time than learning of the directional cues combined with a reduced attention to the impression of distance.

A possible effect of the reduced reverberation time during recordings was studied in experiment Ay. Table 3 compares results from this experiment and from the late real-life experiment with "normal" reverberation. It is seen that there is no significant difference at all. This is not unexpected, since the difference in reverberation time was very small.

In conclusion of the analysis of these possible sources of errors it is accepted that the reverberation time did not influence the results. It is further recognized that, for real life, the number of median-plane errors has decreased and the number of distance errors has increased during the course of the study. In order to arrive at the most conservative conclusions about artificial-head recordings, further statistical comparisons with real life will use the initial value from experiment A for median-plane errors (and out-of-cone and within-cone errors), and a "long-term" value for distance errors obtained as the pooled value of experiments Ax and Ay.

The result of a such modified comparison of distance errors is shown in Fig. 5, and no significant deterioration of performance is seen here for any of the artificial heads. The indications of significance for distance errors in Table 1 and Fig. 4 were put in parentheses in anticipation of this result.

2.4 Recordings at Blocked Ear Canal

The results from recordings at the blocked ear canal are given for the four error categories in Fig. 6, which also includes a comparison with real life. The most noticeable observations are 1) an unmistakable increase in the number of median-plane errors for all heads (signifi-

Table 1. Comparison of real-life performance (experiment A less 12 subjects who did not participate in experiment B) and performance with recordings from artificial heads (recordings with built-in microphones, experiment B). Errors are given in percentage and numbers. Number of stimuli that can result in errors in a category are given in parentheses.[a]

Condition	Error: Out of Cone	Error: Within Cone	Error: Median Plane	Error: Distance	Total Number of Stimuli
Real life	0.2% 2 (912)	0.3% 1 (336)	16.0% 77 (480)	11.9% 40 (336)	912
KEMAR 2	1.3%** 12 (912)	3.6%** 12 (336)	38.5%*** 185 (480)	19.6%(**) 66 (336)	912
KU80i	1.3%** 12 (912)	21.7%*** 73 (336)	50.2%*** 241 (480)	20.8%(**) 70 (336)	912
KU81i	1.5%** 14 (912)	13.7%*** 46 (336)	44.0%*** 211 (480)	22.9%(***) 77 (336)	912
HMS I	2.7%*** 25 (912)	0.9% 3 (336)	37.3%*** 179 (480)	16.1% 54 (336)	912
HMS II	1.9%*** 17 (912)	3.6%** 12 (336)	46.7%*** 224 (480)	12.5% 42 (336)	912
4128	1.2%* 11 (912)	8.9%*** 30 (336)	42.1%*** 202 (480)	17.0%(*) 57 (336)	912
5930	1.1%* 10 (912)	3.9%*** 13 (336)	48.5%*** 233 (480)	16.1% 54 (336)	912
Toronto	0.8% 7 (912)	3.0%** 10 (336)	41.2%*** 198 (480)	17.6%(*) 59 (336)	912

[a] Statistical tests compared each of the artificial heads with real life (one-sided Fisher–Irwin test assuming best performance in real life). Significant *** at 0.1% level, ** at 1% level, and * at 5% level. (See text, especially Section 2.3, for significance levels given in parentheses.)

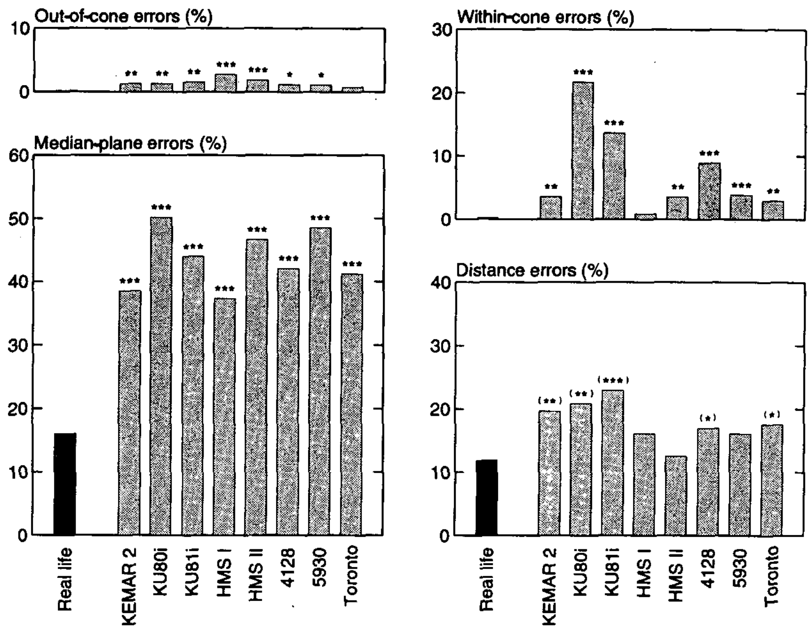

Fig. 4. Comparison of real-life performance (experiment A less 12 subjects who did not participate in experiment B, 912 stimuli) and performance with recordings from artificial heads (recordings with built-in microphones, experiment B, 912 stimuli for each head). Statistical tests compared each of the artificial heads with real life (one-sided Fisher–Irwin assuming best performance in real life). Significance *** at 0.1% level, ** at 1% level, and * at 5% level. (See text, especially Section 2.3, for significance levels given in parentheses.)

cant at the 0.1% level), 2) some increase of out-of-cone errors for most heads (significant at various levels), 3) some increase of within-cone errors for most heads (significant at various levels), KU80i and KU81i being remarkably different, though, in showing very much increased values, and 4) no difference for distance errors. Detailed results for each head are given in Section 2.6.

2.5 Comparison of Recording Types

The significant deviations from real life were in general very similar for recordings with built-in microphones and with blocked ear canal microphones. It would therefore be natural to make a comparison of the results for the two recording types. The result of this analysis is shown in Fig. 7. It is seen that results are very similar for the two situations. Only two comparisons turned out to indicate a significant difference, a significance which we—taking into account the low significance level and the large number of comparisons—consider a coincidence. Hence we conclude that the recording point does not influence the localization performance when proper equalization is accomplished.

2.6 Detailed Results for Each Artificial Head

For each artificial head, information on responses to single source positions may be given for each of the recording types. As we have seen that there is no significant difference between the two types, we have, in order to save space, chosen to present results from only one recording type. Since the experiment with recordings at the blocked ear canal, experiment C, included more subjects and thus a wider range of "fits" between human and artificial heads, we selected the results from this experiment for presentation. The diagrams for individual heads are given in Figs. 8–12.

In general the errors seen are similar to those mentioned in Section 2.2. Only a couple of observations specific for certain heads should be mentioned at this point. For KU80i and KU81i sound sources in the directions LEFT LOW and RIGHT LOW are often perceived in the directions LEFT HIGH and RIGHT HIGH, respectively, whereas these sources are almost always perceived cor-

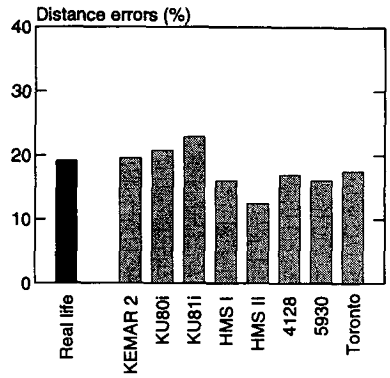

Fig. 5. Comparison of real-life distance performance (experiments Ax and Ay, 1824 stimuli) and performance with recordings from artificial heads (recordings with built-in microphones, experiment B, 912 stimuli for each head). Statistical tests did not show significant difference between any of the artificial heads and real life (one-sided Fisher–Irwin test at 5% significance level assuming best performance in real life).

Table 2. Comparison of real-life performance in original experiment (experiment A) and at a later stage in course of study (experiment Ax). Errors are given in percentage and numbers. Number of stimuli that can result in errors in a category are given in parentheses.[a]

	Error				
Condition	Out of Cone	Within Cone	Median Plane	Distance	Total Number of Stimuli
Real life (experiment A)	0.2% 2 (912)	0.3% 1 (336)	16.0% 77 (480)	11.9% 40 (336)	912
Real life (experiment Ax)	0.0% 0 (912)	0.0% 0 (336)	8.3%*** 40 (480)	22.0%*** 74 (336)	912

*** Significance at 0.1% level in two-sided Fisher–Irwin test.

Table 3. Comparison of real-life performance with different reverberation times (experiments Ax and Ay), both observed at a late stage in course of study. Errors are given in percentage and numbers. Number of stimuli that can result in errors in a category are given in parentheses.[a]

	Error				
Condition	Out of Cone	Within Cone	Median Plane	Distance	Total Number of Stimuli
Real life (experiment Ax)	0.0% 0 (912)	0.0% 0 (336)	8.4% 40 (480)	22.0% 74 (336)	912
Real life (experiment Ay)	0.3% 3 (912)	0.0% 0 (336)	9.4% 45 (480)	16.4% 55 (336)	912

[a] Statistical tests did not show any significant difference between the two conditions (two-sided Fisher–Irwin test at 5% level).

rectly for the other heads. It is worth noting that the KU80i and KU81i are the only heads without a torso. It would be reasonable to assume that lack of shadowing from the shoulders is responsible for errors of this kind.

The HMS I has also problems with these directions, since responses are often given at LEFT and RIGHT. Unpublished measurements of interaural time delays (ITDs) carried out in our laboratory have shown that the ITDs of HMS I are higher than for most humans as well as for the other artificial heads. Thus the ITD of HMS I for the LOW and HIGH directions at the sides may be closer to human ITDs for the directions LEFT and RIGHT.

2.7 Comparison with Recordings from Human Subjects

It is characteristic for artificial-head recordings that they are not made in the ears of the individual listener. We showed earlier that this leads to reduced localization

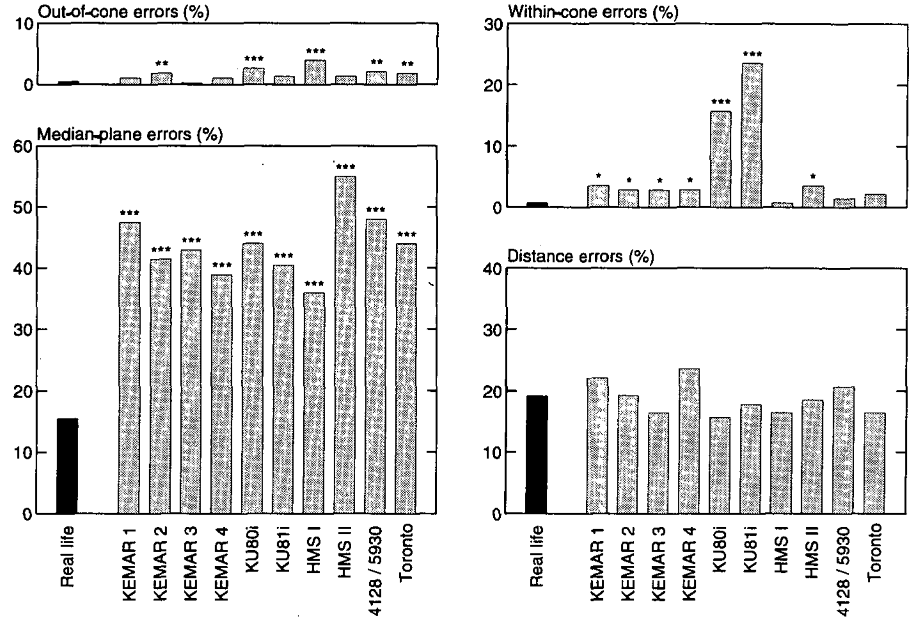

Fig. 6. Comparison of real-life performance (experiment A, 2280 stimuli) and performance with recordings from artificial heads (recordings with blocked ear canal microphones, experiment C, 380 stimuli for each head). Statistical tests compared each of the artificial heads with real life (one-sided Fisher–Irwin assuming best performance in real life). Significance *** at 0.1% level, ** at 1% level, and * at 5% level. (The comparison of distance errors was made with pooled experiments Ax and Ay, 1824 stimuli, see Section 2.3.)

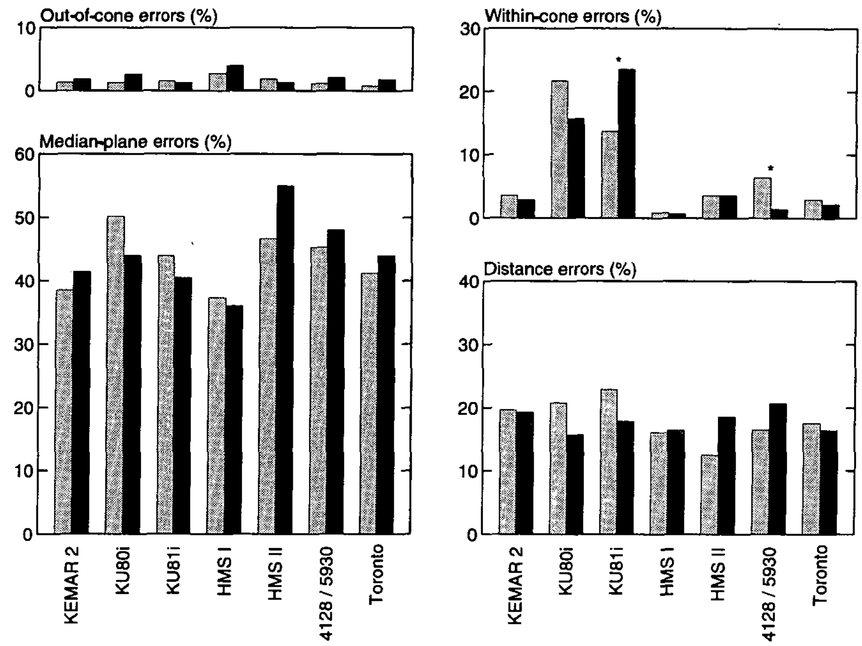

Fig. 7. Comparison of recording types. Light columns—built-in microphones (experiment B, 912 stimuli for each head); dark columns—blocked ear canal microphones (experiment C, 380 stimuli for each head). Statistical tests compared the two recording types for each of the artificial heads (two-sided Fisher–Irwin tests). * Significance at 5% level.

performance [19]. It is therefore interesting to compare the results with information on performance with nonindividual recordings from humans. In Fig. 13 the results for the artificial-head recordings are compared with the results obtained in a previous investigation for listening to recordings made in the ears of other human subjects chosen at random [20].

The localization performance for the artificial heads is seen to be in the same order of magnitude as for recordings with random human subjects. Some artificial heads, though, show more median-plane errors. The many within-cone errors for the KU80i and KU81i still represent exceptions, as does the relatively high number of out-of-cone errors for the HMS I.

We also showed earlier that it is possible to obtain a significantly better localization performance if the nonindividual recordings originate from a carefully selected human recording head [20].

The fact that recordings from a selected human offers such an improvement is due to the relatively large difference between the "best" and the "worst" human recording heads. In Fig. 14 the results from artificial heads (present investigation) are displayed into the ranking of results from human recording heads (our previous investigation [20]). As it can be seen, the artificial heads are found in the poorer part of the human heads with 60% (18) of the human heads being better than even the best artificial head.

2.8 Analysis of Variance

Strictly speaking, the analyses described in the previous subsections are valid only for the particular group of subjects who participated in the experiment. In order to evaluate the results with regard to their general validity for a population—in particular when considering differences between artificial heads—analyses of variance were carried out with subject as random factor. Real-life data were not included in these analyses.

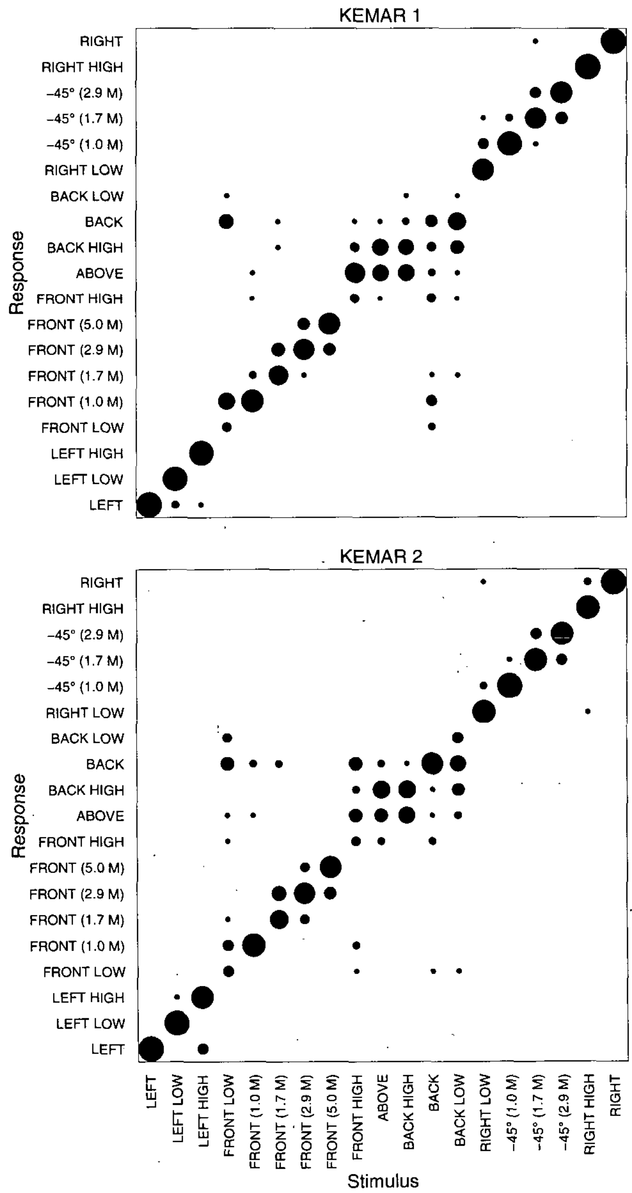

Fig. 8. KEMAR 1 and KEMAR 2 with blocked ear canal microphones (380 stimuli from experiment C in each frame).

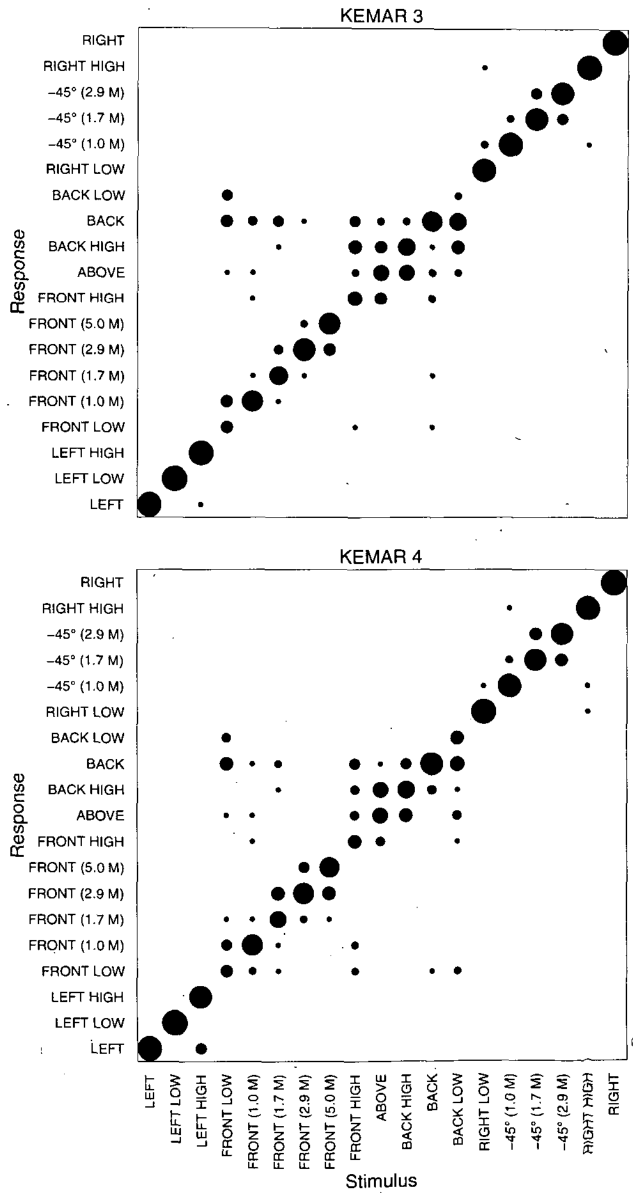

Fig. 9. KEMAR 3 and KEMAR 4 with blocked ear canal microphones (380 stimuli from experiment C in each frame).

A precondition for an analysis of variance is homogeneity of variance. Therefore, Bartlett's tests (as described by Glaser [29]) were carried out initially, and homogeneity of variance was accepted for median-plane and distance errors in both of the experiments B and C, whereas, for both experiments, it was rejected at 5% significance level for out-of-cone and within-cone errors. The rejection for these two error categories is not surprising, since for some recording heads the number of these errors is very small, thus the variance is presumably smaller than for recording heads with more errors. Also in the case of a small number of errors, it is hardly reasonable to assume a normal distribution, another precondition for the analysis of variance.

The analyses of variance are given in Table 4. A significant effect of the recording head is seen for median-plane errors in experiments B and C, whereas no effect is seen for distance errors in any of the experiments.

Fig. 15 shows means and 84% confidence intervals for each of the heads calculated from experiment C. The confidence intervals have been calculated from the common variances from the analyses of variance including the subject term, thus the confidence interval is general for a population. The "odd" 84% size of the confidence intervals has been chosen such that the means of two heads with nonoverlapping confidence intervals will be significantly different at 5% level in a t-test. (Data from experiment C have been chosen for this presentation, since experiment C included more subjects and thus a wider range of "fits" between humans and artificial heads than experiment B.)

2.9 Further Comments

All the listening experiments conducted, required, for obvious reasons, that the listeners kept their heads still during playback. Although the subjects were asked to keep their heads still, and we further monitored the obey-

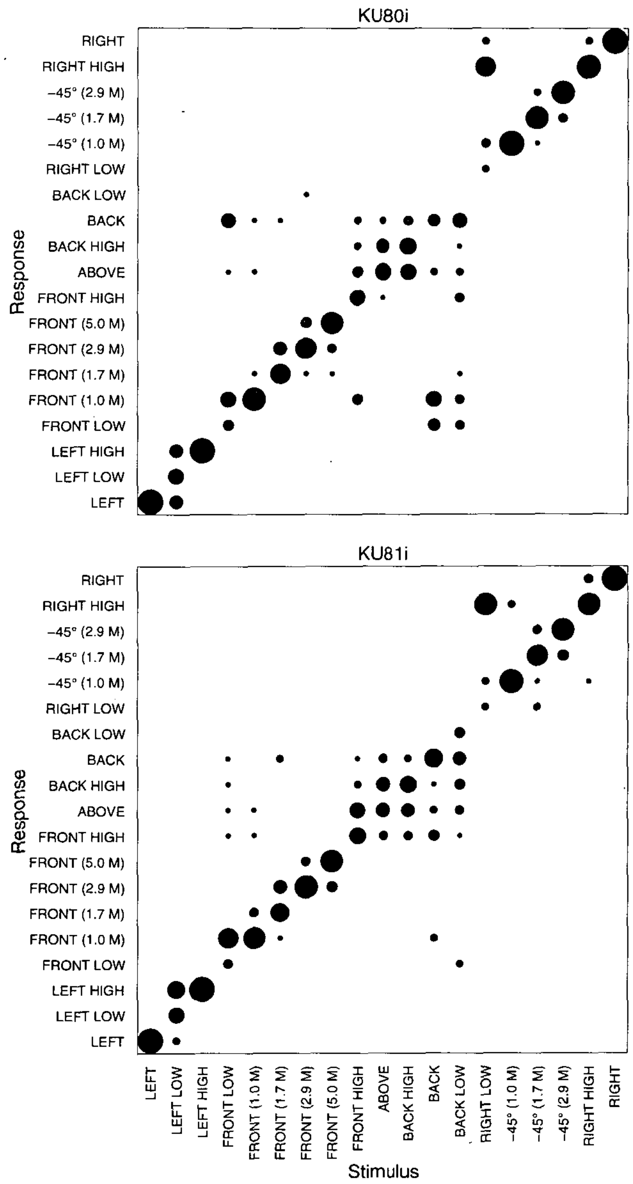

Fig. 10. KU80i and KU81i with blocked ear canal microphones (380 stimuli from experiment C in each frame).

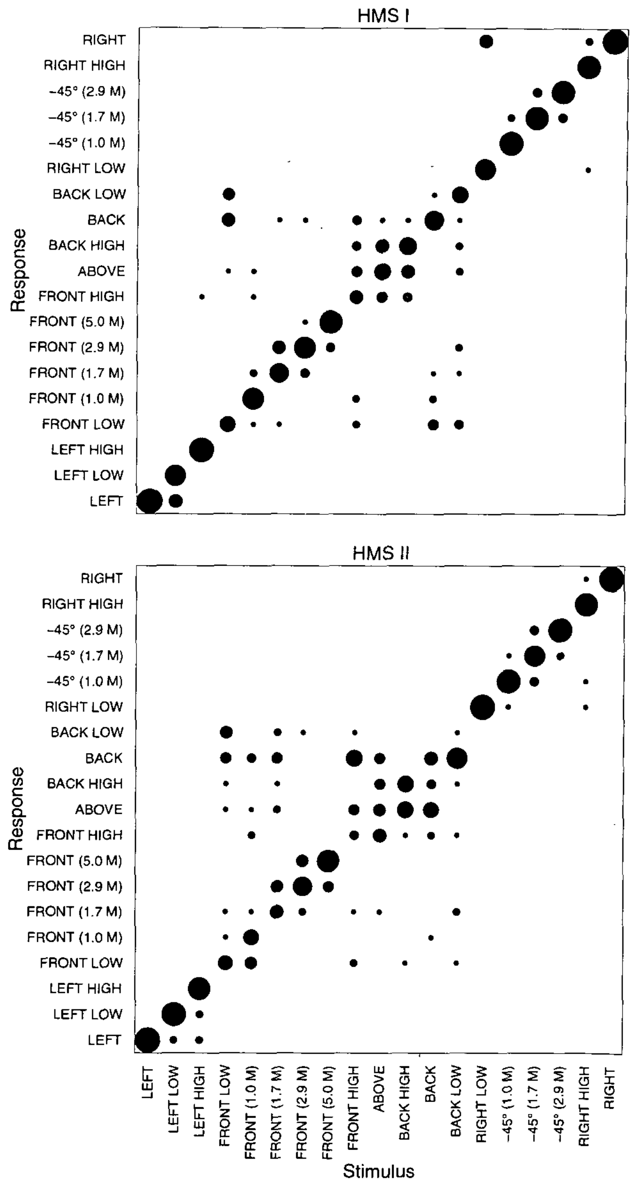

Fig. 11. HMS I and HMS II with blocked ear canal microphones (380 stimuli from experiment C in each frame).

ance of this, it can be argued that the subjects might have induced minor head movements in the real-life situation, which could have assisted the localization and improved the performance. It is therefore important to notice that the same experimental procedure was used in our earlier experiments for the reproduction of individual binaural recordings [19], in which case the localization performance was the same as in the real-life situation. There is thus no indictment that the localization performance in the real-life situation was improved by the use of head movements.

3 CONCLUSION

We have shown that artificial-head recordings do not result in the same localization as observed in real life. This is the conclusion for all artificial heads that were included in the investigation, and it proved valid both for recordings that were made by the artificial heads' built-in microphones, and for recordings made with a common microphone technique (blocked ear canal recordings).

The deteriorated localization performance was reflected in the increase in directional errors. Especially the number of confusions between the sources in the median plane increased greatly (from 15.5% in real life to 36.0–55.0% for the artificial heads).

It is known that nonindividual recordings in general result in more errors in judgment of directions, and the corresponding figure for nonindividual recordings from random human subjects is 36.3% (from a previous investigation [20]). The performance with artificial-head recordings is thus either comparable to the performance obtained from recordings with random humans, or poorer.

The artificial heads could, however, be expected to better represent the "typical" human subject, since the design goal for artificial heads ought to be the geometry that matches the average listener best. In [20] it was found that for the same experimental setup and procedure, a head could be selected which resulted in a better localization performance.

Although it is somewhat disappointing to conclude that the artificial heads at best can be compared to performance obtained with recordings from random human subjects, it is worthwhile to stress the superior reproduction that is obtained from binaural recordings, compared to other reproduction principles in general. It is further encouraging that it was possible to find a human for which recordings gave a better localization performance than for random human subjects or artificial heads. This ought to encourage the design and production of better artificial heads. A study of a number of more recent developments of artificial heads is taking place at our laboratory currently.

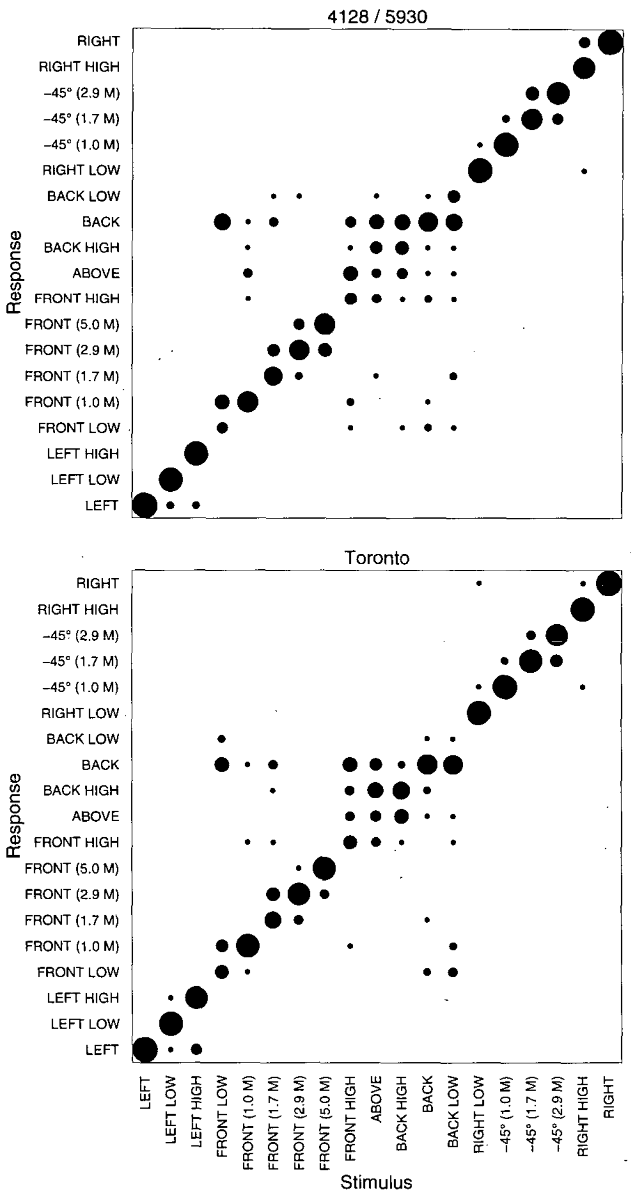

Fig. 12. 4128/5930 and Toronto with blocked ear canal microphones (380 stimuli from experiment C in each frame).

4 ACKNOWLEDGMENT

This work was supported by Brüel & Kjær A/S, Perceptive Acoustics A/S, the National Agency of Industry and Trade, Denmark, the Danish Technical Research Council, and Aalborg University. The authors are further in dept to a number of persons who have made this investigation possible. First of all, we wish to thank manufacturers and institutions who have made the artificial heads available for the experiments. These are Knowles Electronics (for lending us the KEMAR with four different sets of ears), Danish Broadcasting (for lending us the KU80i), Institut für Rundfunktechnik in Munich (for lending us the KU81i), Volvo and Prof. Mendel Kleiner from Chalmers University of Technology in Gothenburg (for lending us the HMS I and bringing it to Aalborg), Head Acoustics GmbH (for lending us the HMS II and providing equalizer data for HMS II, HPS II, and HPS III), Brüel & Kjær (for lending us the 5930), and finally Prof. Hans Kunov from the University of Toronto (for lending us the head called Toronto).

We further wish to thank our former colleagues Kim Alan Larsen and Jørn Vagn Hundebøll for their participation when the experiments were planned and carried out. We also wish to thank Anne Kirstine Andersen for her handling of all appointments with the subjects who took part in the experiments. Also Claus Vestergaard Skipper, the technical employee at this laboratory, deserves appreciation for his valuable help with the practical work. Finally, we wish to thank the 20 persons who came to our laboratory many times and listened patiently to the playback of the speech segment.

5 REFERENCES

[1] P. Damaske and B. Wagener, "Richtungshörversuche über einen nachgebildeten Kopf (Directional Hearing Tests by the Aid of a Dummy Head)," *Acustica*, vol. 21, pp. 30–35 (1969).

[2] P. Damaske and V. Mellert, "Ein Verfahren zur richtungstreuen Schallabbildung des orberen Halbraumes über zwei Lautsprecher (Sound Reproduction of

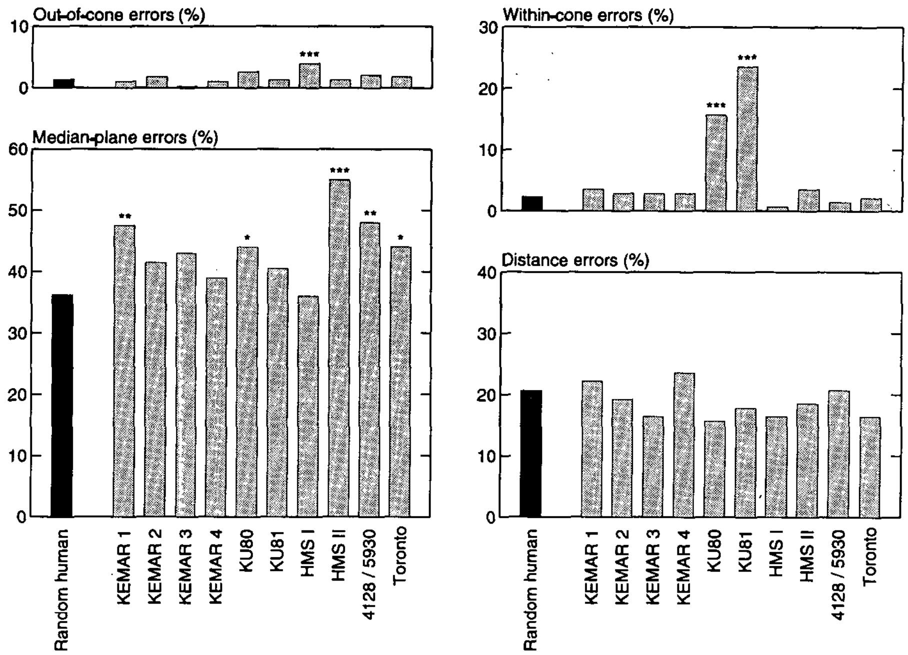

Fig. 13. Comparison of performance with recordings from artificial heads (recordings with blocked ear canal microphones, experiment C, 380 stimuli for each head) and earlier observed performance with recordings from random human subjects (experiment B of [20], 11020 stimuli). Statistical tests compared each of the artificial heads with random human subjects (two-sided Fisher–Irwin tests). Significance *** at 0.1% level, ** at 1% level, and * at 5% level.

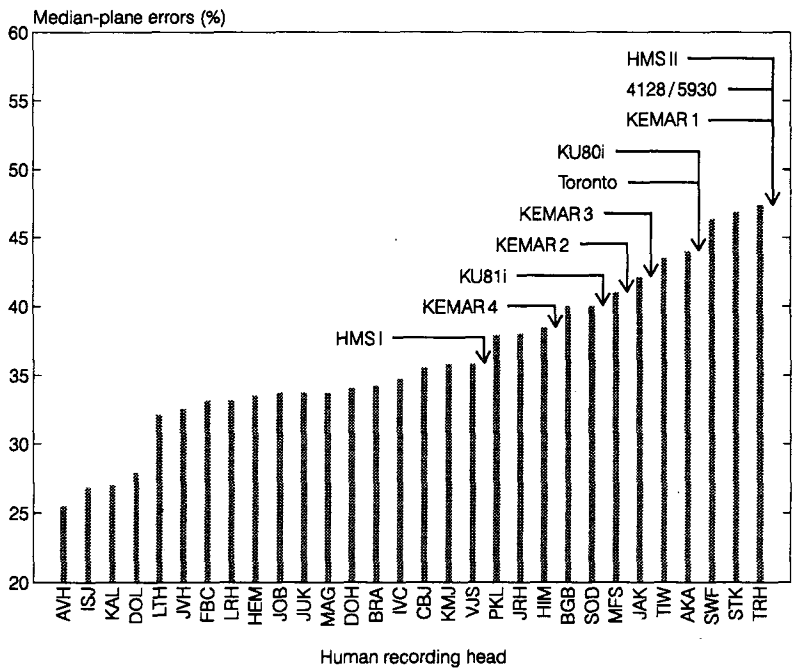

Fig. 14. Median-plane errors of human and artificial recording heads (experiment B of [20] and experiment C of present investigation, respectively). Recording heads are ordered according to error percentage.

the Upper Semi-Space with Directional Fidelity Using Two Loudspeakers)," *Acustica*, vol. 22, pp. 153–162 (1969/70).

[3] H. Wilkins, "Kopfbezügliche Stereophonie—ein Hilfsmittel für Vergleich und Beurteilung verschiedener Raumeindrücke (Head-Related Stereophony—An Aid for the Comparison and Critical Examination of Different Room Effects)," *Acustica*, vol. 26, pp. 213–221 (1972).

[4] P. Laws and H.-J. Platte, "Spezielle Experimente zur kopfbezogenen Stereophonie," in *Fortschritte der Akustik* (DAGA '75), pp. 365–368 (1975).

[5] H.-J. Platte and P. Laws, "Die Vorne-Ortung bei der kopfbezogenen Stereophonie," *Radio Mentor Elektronik*, vol. 42, no. 3, pp. 97–100 (1976).

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

[7] G. Boerger, P. Laws, and J. Blauert, "Stereophone Kopfhörerwiedergabe mit Steuerung bestimmter Übertragungsfaktoren durch Kopfdrehbewegungen (Stereophonic Reproduction by Earphones with Control of Special Transfer Functions through Head Movements)," *Acustica*, vol. 39, pp. 22–26 (1977).

[8] T. Poulsen, "Hörvergleich unterschiedlicher kunstkopfsysteme," *Rundfunktechnisches Mitteilungen*, vol. 22, no. 4, pp. 211–214 (1978).

[9] K. Genuit and H.-J. Platte, "Untersuchungen zur Realisation einer richtungstreuen Übertragung mit elektroakustischen Mitteln," *Fortschritte der Akustik* (DAGA '81), pp. 629–632 (1981).

[10] P. Laws and H.-J. Platte, "Ein spezielles Konzept zur Realisierung eines Kunstkopfes für die kopfbezogene Stereofone Aufnahmetechnik (The Design of a Special Artificial Head for Head-Related Stereophony)," *Rundfunktechnisches Mitteilungen*, vol. 22, no. 1, pp. 28–31 (1978).

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

[12] G. Theile and G. Spikofski, "Vergleich zweier Kunstkopfsysteme unter berücksichtung Verschiedener Anwendungsbereiche (Comparison Between Two Dummy Head Systems with Due Regard to Different Fields of Application)," *Fortschritte der Akustik* (DAGA '84), pp. 223–226 (1984).

[13] K. Genuit, "Simulation des Freifeldes über Kopfhörer zur Untersuchung des räumlichen Hörens und der

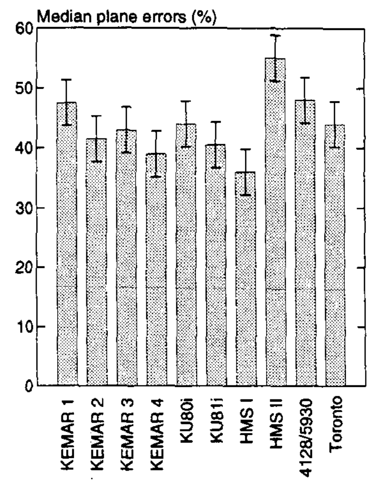

Fig. 15. Median-plane errors with recordings from artificial heads (experiment C). Columns indicate observed means, bars 84% confidence intervals. The means of two heads with non-overlapping confidence intervals will be significantly different at 5% level in a *t*-test.

Table 4. Analyses of variance for error percentages (repeated measures, subject as random factor).

Source	SS	df	MS	F	p
Experiment B, median-plane errors					
Recording head	1206.6	7	172.4	2.57	0.0245
Subject	2305.1	7	329.3		
Recording head × subject	3287.5	49	67.1		
Total	6799.2	63			
Experiment B, distance errors					
Recording head	588.8	7	84.1	1.24	n.s.
Subject	5916.0	7	845.1		
Recording head × subject	3336.2	49	68.1		
Total	9841.0	63			
Experiment C, median-plane errors					
Recording head	5150.5	9	572.3	4.75	<0.001
Subject	6985.5	19	367.7		
Recording head × subject	20599.5	171	120.5		
Total	32735.5	199			
Experiment C, distance errors					
Recording head	1302.0	9	144.7	1.09	n.s.
Subject	37709.0	19	1984.7		
Recording head × subject	22799.5	171	133.2		
Total	61791.4	199			

Sprachverständlichkeit (Simulation of the Freefield Using Headphones for Investigation of the Spatial Hearing and Speech Intelligibility)," *Audiologische Akustik*, vol. 6, pp. 202–221 (1988).

[14] G. Krumbacher, "Über die Leistungsfähigkeit der kopfbezüglichen Stereofonie (On the Efficiency of Head-Oriented Stereophony)," *Acustica*, vol. 21, pp. 288–293 (1969).

[15] G. Plenge, "Über das Problem der Im-Kopf-Lokalisation (On the Problem of In Head Localization)," *Acustica*, vol. 26, no. 5, pp. 241–252 (1972).

[16] R. G. Wettshurek, "Die absoluten Unterschiedsschwellen der Richtungswarhnehmung in der Medianebene beim natürlichen Hören, sowie beim Hören über ein Kunstkopf-Übertragungssystem (The Absolute Difference Limen of Directional Perception in the Median Plane under Conditions of Both Natural Hearing and Hearing with Artificial-Head-System)," *Acustica*, vol. 28, no. 4, pp. 197–208 (1973).

[17] G. Plenge, "On the Differences Between Localization and Lateralization," *J. Acoust. Soc. Am.*, vol. 56, no. 3, pp. 944–951 (1974).

[18] F. E. Toole, "Binaural Record/Reproduction Systems and Their Use in Psychoacoustic Investigations," presented at the 91st Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 39, p. 1005 (1991 Dec.), preprint 3179.

[19] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural Technique: Do We Need Individual Recordings?," *J. Audio Eng. Soc.*, vol. 44, pp. 451–469 (1996 June).

[20] H. Møller, C. B. Jensen, D. Hammershøi, and M. F. Sørensen, "Using a Typical Human Subject for Binaural Recording," presented at the 100th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 44, p. 632 (1996 July/Aug.), preprint 4157, pp. 1–18. Final report in preparation for submission to *J. Audio Eng. Soc.*

[21] IEC 711, "Occluded-Ear Simulator of the Measurement of Earphones Coupled to the Ear by Ear Inserts," 1st ed., International Electrotechnical Commission, Geneva, Switzerland (1981).

[22] ANSI S3.25-1979 (R1986), "American National Standard for an Occluded Ear Simulator," American National Standards Institute, New York (1980).

[23] H. Møller, "Fundamentals of Binaural Technology," *Applied Acoustics*, vol. 36, no. 3/4, pp. 171–218 (1992).

[24] D. Hammershøi and H. Møller, "Sound Transmission to and within the Human Ear Canal," *J. Acoust. Soc. Am.*, vol. 100, pp. 408–427 (1996 July).

[25] IEC 268-13, "Sound System Equipment, Pt. 13, Listening Tests and Loudspeakers," International Electrotechnical Commission, Geneva, Switzerland (1985).

[26] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-Related Transfer Functions of Human Subjects," *J. Audio Eng. Soc.*, vol. 43, pp. 300–321 (1995 May).

[27] 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. 203–217 (1995 Apr.).

[28] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists* (Wiley, New York, 1987), pp. 230–234.

[29] R. E. Glaser, in Samuel Kotz and Norman L. Johnson (Eds.), "Encyclopedia of Statistical Sciences," vol. 1, Wiley & Sons, New York, pp. 189–191.

THE AUTHORS

H. Møller

D. Hammershøi

C. B. Jensen

M. F. Sørensen

Henrik Møller was born in Århus in 1951. He studied electrical engineering (Danish Engineering Academy) and received a B.Sc. degree in 1974. He worked as a development engineer for Brüel & Kjær from 1974 to 1976. Since then he has been at Aalborg University. He became associate professor in 1980, received a Ph.D. degree in 1984, and was appointed professor in 1988. During the period 1991–94 he was partly on leave from the university to work as a director of Perceptive Acoustics A/S, a research subsidiary company of Brüel & Kjær.

Dr. Møller's previous and current research reflect his long-time experience with sound, its influence on humans, acoustical measurement techniques, signal processing, hearing, and psychometric methods. His research areas include effects of infrasound and low-frequency noise on humans, investigations of hearing thresholds and loudness assessment, and exploitation of binaural techniques. He is the author of numerous scientific publications and invited as well as contributed conference papers.

When new high-quality acoustical laboratories were built at Aalborg University in 1987, Dr. Møller was responsible for the design as well as control of the work. As head of the Acoustics Laboratory, he is now the manager of research and education in a wide range of areas such as human sound perception, audiology, psychometry, electroacoustics, recording and playback techniques, auralization in acoustic room modeling and virtual reality, acoustical measurement techniques, electronics, and signal processing.

Dr. Møller has organized conferences on Low Frequency Noise and Hearing (Aalborg 1980) and general acoustics (Nordic Acoustical Meeting, Aalborg 1986). He is convener of ISO Technical Committee 43: "Acoustics," Working Group 1: "Thresholds of Hearing," a member of Working Group 6: "Determination of Noise Immissions from Sound Sources Placed Close to the Ears," and of the editorial board of Journal of Low Frequency Noise & Vibration. He holds membership in the Danish Engineering Society, Audio Engineering Society, Acoustical Society of America, IEEE, Danish Acoustical Society, Danish Technical-Audiological Society, Danish Biomedical Society, and Danish Standardisation Organisation (board of Acoustics and Working Group of Audiometry and Hearing).

Dr. Møller spends hours off work (too few) by playing big band music on his baritone saxophone or by keeping his classic British cars in good shape. Now and then, he also drives them.

•

Dorte Hammershøi was born in 1965. She studied electrical engineering in Aalborg (Aalborg University). In 1989 she received a Master of Science in electrical engineering, with specialization in biomedical engineering. During her study, she worked part time as an engineer at Synaps Electronic Aps., developing communication aids for motionally disabled people.

From 1990 to 1995 she worked as a research engineer at the Acoustics Laboratory at Aalborg University, and in 1995 she received the Ph.D. degree in acoustics. Since 1995 she has worked as an assistant professor at Aalborg University.

Dr. Hammershøi is experienced in electronics, sound measuring techniques, digital signal processing, hearing, and psychometry, and is familiar with neuro and sensory physiology. From 1990 to 1992 she worked on research projects on the improvement of sound reproduction techniques by means of binaural techniques. Since 1992, she has been working with computer generation of binaural signals for auralization of room models and for generation of auditory virtual environments in virtual reality systems.

Dr. Hammershøi holds membership of the ISO Technical Committee: "Acoustics," Working Group 6: "Determination of Noise Immissions from Sound Sources Placed Close to the Ears," the Danish Acoustical Society, Danish Engineering Society, and Audio Engineering Society, and is the author of many scientific papers.

•

Clemen Boje Jensen was born in Aalborg, Denmark, in 1965. He studied electrical engineering at Aalborg University. In 1991 he received a Master of Science in electrical engineering, specializing in acoustics.

From 1991 to 1994 he was employed as a research engineer at the Acoustics Laboratory of Aalborg University, where he worked on various aspects of binaural techniques, including sound measurements in the human ear canal, design of articial heads, calibrated binaural sound reproduction, design of measuring and experimental setups, and artificial generation of binaural signals by means of digital signal processing. Since 1995 he has been working in the R&D department of Bosch Telecom Danmark A/S (formerly Dancall Telecom A/S), developing DECT telephones.

Mr. Jensen holds membership in the Danish Acoustical Society, Danish Engineering Society, and the Audio Engineering Society.

•

Michael Friis Sørensen was born in 1966. He studied electrical engineering at Aalborg University, where he received his M.Sc.E.E., specializing in acoustics, in 1991. From 1991 to 1995 he worked as a research engineer at Aalborg University.

His major fields of interest are psychoacoustics, noise measuring techniques, psychometry, electroacoustics, audiology, and experimental design. From 1991 he has worked on research projects within various aspects of binaural technique and audiometry. Since 1995 August he has been working as a consultant in noise and vibration control at Ødegaard & Danneskiold-Samsøe, Jylland Aps.

Mr. Sørensen is a member of the Danish Acoustical Society, Danish Engineering Society, Audio Engineering Society, and Danish Technical Audiological Society.