Proceedings of the Institute of Acoustics

STEREOPHONIC LOCALISATION IN ROOMS, COMPARING THE DISTRIBUTED-MODE LOUDSPEAKER (DML) WITH CONVENTIONAL TWO-WAY CONE-BASED LOUDSPEAKERS.

Neil Harris, Sheila Flanagan

New Transducers Ltd (NXT), Huntingdon, UK

1. INTRODUCTION

The subjective effects of the listening environment on listening tests are well documented [1], [2], [3], [4]. In order to avoid these variations, image localisation tests are usually carried out under near anechoic conditions, with a single, well-defined listener location. In normal use, however, loudspeakers are listened to in rooms with reflective boundaries, and by more than one person at a time. Conventional loudspeakers, as specular acoustic radiators, are particularly susceptible to acoustic room modes and boundary interference, and have a limited "sweet-spot" for optimal listening. In contrast, diffuse acoustic radiators such as the Distributed Mode Loudspeaker (**DML**) [5], are cited as having "sympathetic boundary interactions" and no sweet-spot [6]. The nature and acoustical properties of DMLs are discussed elsewhere, e.g. [7], [8], but for the purposes of this paper they may be considered as having a very wide directivity and a largely diffuse radiation.

A hypothesis is proposed that diffuse acoustic radiators, such as distributed-mode loudspeakers (DML), lessen the degradation caused by room acoustics on stereophonic localisation. A series of double blind psychometric listening tests was proposed to compare the ability of these two classes of radiator to localise a stereo image in an untreated room. Results from two such series of tests were presented at the 104th and 105th AES conventions in Amsterdam and San Francisco [9][10]. In contrast to anecdotal evidence, the results showed that the performance of the conventional loudspeaker system was worse than the DML system, even when tested from the "sweet spot".

2. BACKGROUND – HOW DO WE LOCALISE AN IMAGE?

For a full description of the processes involved, the reader is referred to published works on the subject, for example [11]. Localisation depends on the *dichotic* properties of sound – i.e. on differences between the sound reaching the two ears. A synopsis of these properties is included below.

The first important work on human sound localisation was done by Venturi and appeared between 1796 and 1801. From his observations, Venturi advanced an intensity hypothesis, which stated that the difference in simultaneous intensity at the two ears informs us of the true direction of sounds. It was about seventy years later, in 1877, when Rayleigh verified the experiment, and showed that the binaural intensity ratio was a cue for localisation. He also suggested that the strength of this cue was a function of frequency, being strongest at high frequency tones, and negligible for tones whose wavelength was longer than four times the circumference of the head. Many modern texts describe the intensity ratio as an inter-aural level difference or ILD. Some writers use loudness instead of level.

In 1907, Rayleigh went on to demonstrate that at low frequencies, where ILDs are small, phase becomes an important cue to localisation. Later authors either confirmed this, or suggested time differences as an alternative. The modern measure used is the inter-aural time difference or ITD, as this is applicable to tonal and non-tonal sounds. In the case of the average human head, the maximum usable frequency for ITD cues is about 1500 Hz.

With both ILD and ITD, it is possible to define a "cone of confusion", which has its axis through the ears. All sources on this cone will produce equal ILD and ITD values, and additional cues are required to discriminate between them. The outer ear, or pinnae, is thought to play a significant role in the localisation of sound in the vertical plane, and front-back differentiation. The pinnae only affect high frequencies, above about 3kHz.

When the sound source is in a room, sound arrives at the ear from a variety of directions. Reflected sounds travel longer distances, and by multiple routes, so arrive later and from many directions. Fortunately, localisation seems to be based on the sound that arrives first – the precedence effect – although it is possible to defeat this mechanism under some circumstances. See also [12][13][14][15].

3. THE TEST PROCEDURE

The tests consisted of a sequential presentation of two bursts of pink noise. Each burst was three seconds long separated by a gap of one second. The bursts were of identical origin, only one had been amplitude panned to one of a predefined range of image positions. Subjects were required to identify the burst closer to their right-hand wall[1], and to record this on a response form. They were told that if they were unsure of the answer they were to guess.

The format of the experiment followed the "2 Alternate Forced Choice" method (2AFC). This was a double blind test, following the format recommended by Lipshitz and Vanderkooy [16]. The test signal was band-limited pink noise, 250 Hz to 20 kHz; the random test sequence and

---

[1] In the earlier test format, the subjects were asked to record which of the two signals had been panned away from centre. It was found that the task of determining the absolute central image caused additional errors, owing to room acoustics or loudspeaker matching. By changing the question asked, the localisation task is made differential in nature, and hence it is not necessary to have a known reference signal. Subjects noted that this test procedure was easier than the original.

presentation order being generated in software [17]. The signal format was two channel, dual mono, 44.1 kHz, 16 bit. The replay levels were adjusted to give a Sound Pressure level (SPL) at the subject's position of 70 dBA (Background traffic noise = 40 dBA).

In tests 2 through 5, the offset and presentation order were randomised. This was repeated 32 times. The first six pairs of bursts were predefined in the form of a preamble or teaching session, the remaining 28 being the randomised test signals. Following analysis of these results, the presentation was changed slightly. In tests 9A through 10C, the offset and presentation order were *shuffled* randomly to ensure equal representation of all angles and presentation orders. This was repeated 51 times. The first five pairs of bursts were predefined in the form of a preamble or teaching session, the remaining 46 being the randomised test signals.

Starting with a large displacement angle ϕ, the preamble is there to familiarise the subject with what is required, and to assure them that the test is actually working. The results for the preamble pairs were noted on the response form for the subjects' benefit.

A plan of the room used for the experiments is included as Figure 1, which also indicates the placement of loudspeakers and listeners. The subjects sat in one of three positions, referred to seats A, B and C. Two loudspeaker sources were used and were placed equidistantly, 3 metres from each other and from seat B. Thus the loudspeakers were placed at an angle of 30° either side of the subject B's straight-ahead position.

Five tests were conducted using the "which was panned" paradigm, identified as Test 1 through Test 5, each test representing a different loudspeaker system or placement. Test 1 was used to "try the waters" with a few selected volunteers, and its results are not included in the subsequent analyses. Each subject was tested at seating position B only.

Six further tests, using the "closer to the right" paradigm, were conducted in two groups of three, identified as Test 9 A to C and Test 10 A to C, each test representing a different loudspeaker system or seating arrangement. Each subject was tested at all seating positions.

4. ANECDOTAL EVIDENCE

Feedback from Test 1 allowed the authors to refine the method of recording responses. The authors learnt a lot about human psychology from this process! Originally, the response form had been arranged in the form of columns, but the visual association caused several subjects to confuse yes/no and left/right. As a result, the response form was arranged in the form of rows. One of the biggest problems seemed to be that people don't like being forced to choose between two apparently identical presentations. In addition, many difficulties with the test were due to the fact that engineers hate reading instructions! Consequently, the instructions were read to each participant before each test.

All the tests were amicably received. Many subjects reported that the second set of tests was easier than the first. The consensus of opinion from tests 2 to 5 was that the conventional hi-fi box loudspeakers used in Test 3 produced the tightest image. Many listeners observed that images to the right of centre were easier to assess than those to the left of centre.

As a result of this verbal feedback, the authors were expecting that Test 3 results would stand out from the remainder of the first set, leaving the original hypothesis open until off-axis listening tests are concluded. We were surprised, therefore, to discover that on inspection Test 3 gave the worst results. Our suspicions were partially justified, when a comparison on listener performances showed that listener 10 in Test 3 was seriously at variance with the group average. Test 3a is therefore merely Test 3, with listener 10's results excluded. Identification of each test set-up is found in the results summary given in Figures 10 and 11.

Test 8 (not included here) produced anomalous results, which were eventually tracked down to the cosmetic grilles. These grilles were claimed to be acoustically transparent, but they impaired the ability of the loudspeaker to localise to a statistically significant degree. Anyone involved in industrial design please take note!

5. RESULTS

The results are presented (in a format described later) in Figures 2 to 10. Tests 9 and 10 are given pairwise, thus allowing direct visual comparison of the relative performance of the two pairs of loudspeakers. Figures 11 to 14 summarise the results in tabular form, and indicate the statistical significance of any differences between the loudspeakers. Observation of Figures 7 to 9 suggests that the results of experiment 10 are more precise and more symmetrical off axis than those for experiment 9. Note especially that the localisation error is symmetrical for experiment 10 A and C, indicating the expected steering of the image towards the nearest loudspeaker. In contrast, experiment 9 shows a systematic shift of image owing to interaction between the loudspeakers and the room.

The anecdotal evidence of a bias in the left / right imaging of the room is revealed in the results, with figures 4 and 7 showing it particularly well. In contrast to the anecdotal evidence, however, the results show that the conventional loudspeaker system performed the worst, even when tested from the "sweet spot". Obviously, a small image is not necessarily a well-located image. One might say the cone loudspeaker was lying convincingly, while the DMLs were on average truthful!

6. ANALYSIS OF RESULTS

It is shown in the appendix that the number of correct results for a particular angle and loudspeaker follows the binomial distribution with probability $p_c = (p_r+1)/2$, where p_r is the probability that the loudspeaker can resolve that angle. N was variable in the earlier tests (17 or 34), but equalled 30 for all angles in tests 9 and 10. Figure 4 shows a summary of tests 2 through 5, with 95% confidence limits using the binomial distribution. It shows the average results of all four tests, with each individual result overlaid. A crude ranking of the different loudspeaker systems can be formed by subtracting the number of results that were significantly worse than average from number of results that were significantly better than average, as shown in figure 11. Figure 10 shows a similar summary of tests 9 and 10.

Confidence testing against $p_r = 0$ (i.e. no one can tell, or everyone is guessing) prove almost nothing - nearly all the results are significant at 95%. We therefore propose two tests to determine which side of the *threshold of detection* the results lie.

The threshold of detection is taken to be when $p_r = 0.5$, hence at this threshold we expect 75% successful detection (i.e. $p_c = 0.75$). The two-sided 95% confidence limits are at 3 and 12 wrong answers for N=30. The null hypothesis, H_0, is that we are at the threshold of detection - results above the 97.5% limit are definitely above the threshold; results below the 2.5% limit are definitely below the threshold. The angular resolution of each loudspeaker system by this method is tabulated in Figures 10 and 12.

The second test seeks to equalise type I and type II errors [18]. The three hypotheses being tested are; H_0, we are below threshold of detection, $p_r = 0$; H_1, we are at threshold of detection, $p_r = 0.5$, H_2, we are above threshold of detection, $p_r = 1$. This method of testing results in two critical probabilities, which reject first H_0, then H_1. Figures 2 to 5, and 7 to 9 show the test results scaled so that they cross unity at the critical probabilities. The results which reject H_0 but not H_1 are termed "may hear", and those which reject both H_0 and H_1 are termed "can hear". The angular resolution of each loudspeaker system by this method is also tabulated in Figures 11 and 13.

The total numbers of erroneous results for each loudspeaker pair are compared for each seating position. The statistical significance of the differences is determined and presented in Figures 12 and 14. Do to the greater consistency of the test procedure for tests 9 and 10, the results presented in figure 14 are the more reliable. It is seen that experiment 10 yielded results that were significantly better than those from experiment 9. The level of significance is 95% (one-sided) or better for each seating position, and 99% (one-sided) overall. The significance of improvement is greatest off axis.

7. FUTURE WORK

We would like to have a theoretical model to explain the results presented here. At this time we might speculate on the mechanisms at work, but we are optimistic that a rational and objective explanation is not too far away. Analysis of measurements taken in an anechoic chamber [19] shows that the cross-correlation of left and right ear signals produce different patterns for cones and DMLs. Subsequent analysis of dummy head measurements using a binaural acoustic model [20][21][22] also revealed interesting differences. The details of such analyses will, no doubt, be the subject of future papers.

These results of these experiments have implications on multi-loudspeaker systems and multi-channel audio. Contemporary five or six channel home theatre systems, which aim to present a consistent audio experience to several people simultaneously, may benefit from the use of the new loudspeaker technology. Immersive virtual reality environments, such as that described by Rimell and Hollier [23][24] may also benefit. Extending the research into these areas forms a natural progression to this paper.

8. CONCLUSIONS

A large number of subjective listening tests have been performed. The results from this series of tests confirm the hypothesis that diffuse acoustic radiators, such as distributed-mode loudspeakers (DML), lessen the degradation caused by room acoustics on stereophonic localisation. In fact, the diffuse radiators performed significantly better than a quality two-way cone-in-box loudspeaker system even in the "sweet-spot" which, under anechoic conditions, could be assumed to favour the latter.

9. ACKNOWLEDGEMENTS

The authors would like to thank their colleagues at NXT, and visiting consultants, for their participation in these tests. Their involvement not only made this paper possible, but also allowed us to refine the test procedures over time.

10. APPENDIX

Assume a binomial distribution, with p(able to resolve) = p_r.
Then in N trials, the probability of n correct detections, p_n is given by

$$p_n(n) = \binom{N}{n} \cdot p_r^{\,n} \cdot q_r^{\,N-n}$$

where $q_r = 1 - p_r$, and $\binom{N}{n} = {}_N C_n = \frac{N!}{(N-n)! \cdot n!}$

If we assume that those who do not correctly detect just guess fairly, then 50% of them will get the right answer, i.e. p(guess) = q(guess) = 1/2. So, for n correct detections, the conditional probability of m lucky guesses is

$$p_{nlm}(n,m) = \binom{N-n}{m} \cdot \left(\frac{1}{2}\right)^m \cdot \left(\frac{1}{2}\right)^{N-n-m} = \binom{N-n}{m} \cdot \left(\frac{1}{2}\right)^{N-n}$$

So the probability, p_k, of k correct responses is just the sum of all probabilities where n+m=k, i.e.

$$p_k(k) = \sum_{i=0}^{k} p_n(i) \cdot p_{nlm}(i, k-i)$$

$$p_k(k) = \sum_{i=0}^{k} \binom{N}{i} \cdot p_r^{\,i} \cdot q_r^{\,N-i} \cdot \binom{N-i}{k-i} \cdot \left(\frac{1}{2}\right)^{N-i}$$

$$p_k(k) = \sum_{i=0}^{k} \frac{N!}{(N-i)! \cdot i!} \cdot p_r^{\,i} \cdot q_r^{\,N-i} \cdot \frac{(N-i)!}{(N-k)! \cdot (k-i)!} \cdot \left(\frac{1}{2}\right)^{N-i}$$

$$p_k(k) = \left(\frac{1}{2}\right)^N \cdot \sum_{i=0}^{k} 2^i \cdot \frac{N!}{(N-k)! \cdot (k-i)! \cdot i!} \cdot p_r^{\,i} \cdot q_r^{\,N-i} = \frac{N!}{(N-k)! \cdot k!} \cdot \left(\frac{1-p_r}{2}\right)^{N-k} \cdot \left(\frac{1+p_r}{2}\right)^k$$

It is seen by inspection that we have another binomial distribution, with modified probability p_c given by

$$p_c = \frac{1+p_r}{2} \quad q_c = 1 - p_c = \frac{1-p_r}{2} \quad p_k(k) = \binom{N}{k} \cdot p_c^{\,k} \cdot q_c^{\,N-k}$$

$$E(k) = N \cdot p_c = \frac{N}{2} \cdot (1 + p_r)$$

$$\mathrm{var}(k) = N \cdot p_c \cdot q_c = \frac{N}{4} \cdot \left(1 - p_r^{\,2}\right)$$

11. REFERENCES

[1] M. Barron, "*The Subjective effects of first reflections in concert halls – the need for lateral reflections*", Journal of Sound and Vibration (1971) 15 (4), (475-494).

[2] Sean E. Olive, Peter L. Schuck, Sharon L. Sally and Marc E. Bonneville, "*The Effects of Loudspeaker Placement on Listener Preference Ratings*", JAES vol. 42, No. 9, pp 651-669,

[3] K. R. Holland, P. R. Newell; "*Loudspeakers, Mutual Coupling and Phantom Images in Rooms*", 103rd AES Convention, pre-print # 4581

[4] B. Theiss, M. O. J. Hawksford; "*Loudspeaker Placement for Optimized Phantom Source Reproduction*", 100th AES Convention, pre-print # 4246

[5] A generic term used by New Transducers Limited (NXT) to describe loudspeakers working according to their teaching.

[6] H. Azima & N. Harris, "*Boundary Interaction of Diffuse Field Distributed-Mode Radiators*", 103rd AES Convention, pre-print #4635

[7] N. Harris, "*A Simple Electro-mechanical model of the Distributed Mode Loudspeaker (DML)*", Proceedings of IOA, Vol. 19: Part 6 (1997) – Reproduced Sound 13

[8] G. Bank & N. Harris, "*Distributed Mode Loudspeakers - Theory and Practice*", paper given at AES conference "The Ins and Outs of Audio", in London, England; March 1998

[9] N. Harris, S. Flanagan, M. O. J. Hawksford, "*Stereophonic localization in the presence of boundary reflections, comparing specular and diffuse acoustic radiators.*", 104th AES Convention pre-print #4684

[10] N. Harris, S. Flanagan, M. O. J. Hawksford, "*Stereophonic Localization in Rooms, Comparing Conventional and Distributed-Mode Loudspeakers.*", 105th AES Convention.

[11] W. L. Gulick, G. A. Gescheider, R. D. Frisina, "*Hearing: physiological acoustics, neural coding, and psychoacoustics*", Oxford University Press, 1989. – Chapter 13.

[12] W.M. Hartmann; "*Localization of sound in rooms I*", JASA 74 (5), pp 1380-1391

[13] W.M. Hartmann; "*Localization of sound in rooms II – The effects of a single reflecting surface*", JASA 78 (2), pp 524-533

[14] W.M. Hartmann; "*Localization of sound in rooms III – Onset and duration effects*", JASA 80 (6), pp 1695-1706

[15] W.M. Hartmann; "*Localization of sound in rooms IV – the Fransen effect*", JASA 86 (4), pp 1366-1373

[16] S. P. Lipshitz, J. Vanderkooy, "*The Great Debate: Subjective Evaluation*", JAES Vol.29, No. 7/8, July/August 1981 .

[17] CoolEdit digital audio editor for Windows, by D.Johnston. Syntrillium Software Corporation.

[18] L. Leventhal, "*Type 1 and Type 2 Errors in the Statistical Analysis of Listening Tests*", JAES vol. 34, No. 6, pp437-453

[19] H. Azima, P.Mapp; "*Diffuse Field Distributed Mode Radiators and their Associated Early Reflections*", 104th AES Convention pre-print #4759

[20] V. Pulkki, M. Karjalainen, J. Huopaniemi; "*Analyzing Virtual Sound Attributes using a Binaural Auditory Model*", 104th AES Convention pre-print #4697

[21] W. Lindemann; "*Extension of a binaural cross-correlation model by contralateral inhibition.*", JASA 80 (6) pp1608 ff.

[22] B. Theiss, M. O. J. Hawksford; "*Objective Assessment of Phantom Images in a 3-Dimensional Sound Field using a Virtual Listener*", 102nd AES pre-print #4462

[23] A. Rimell & M. Hollier, "*Reproduction of Spatialised Audio in Immersive Environments with Non-Ideal Acoustic Conditions*", 103rd AES Convention, pre-print #4543.

[24] A. Rimell & M. Hollier, "*Design and Implementation of 3-Dimensional Spatial Audio for Immersive Environments*", Proceedings of IOA, Vol. 19: Part 6 (1997) – Reproduced Sound 13

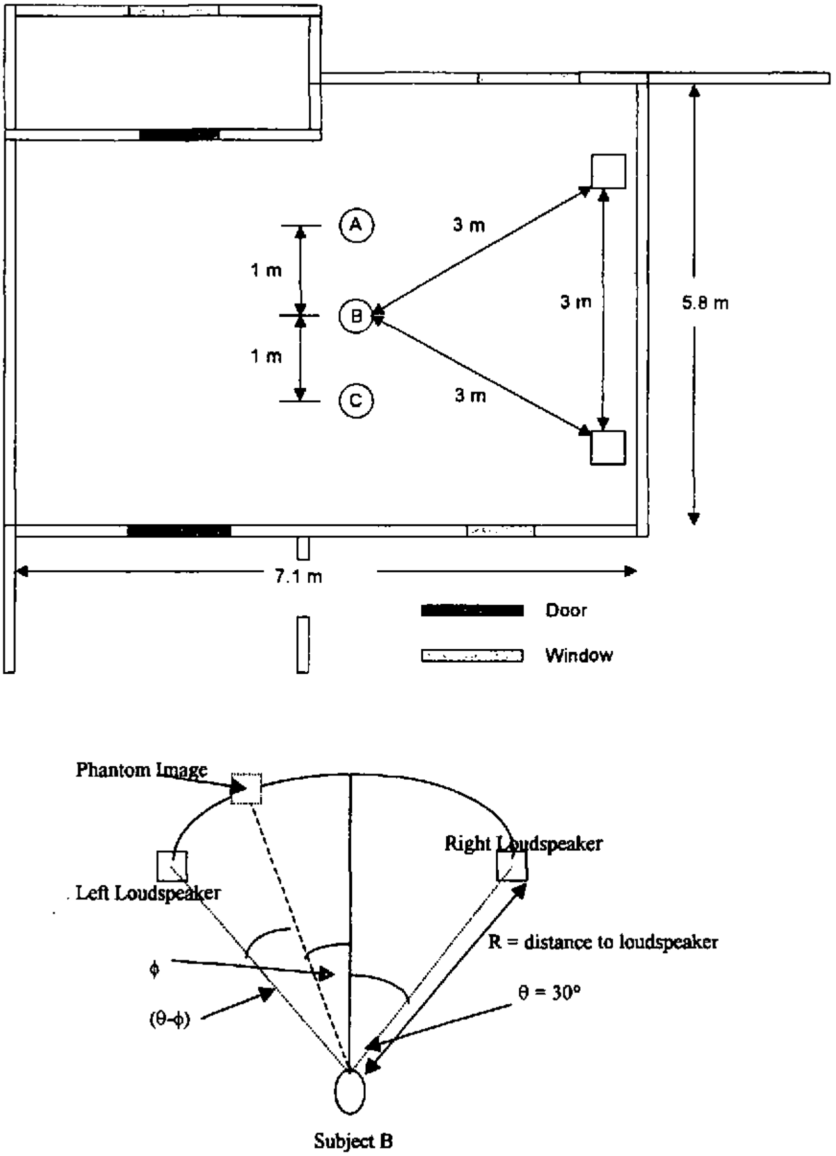


Figure 1. Schematic of the room used for the listening experiments, and the relative positions of source, subject and loudspeakers.

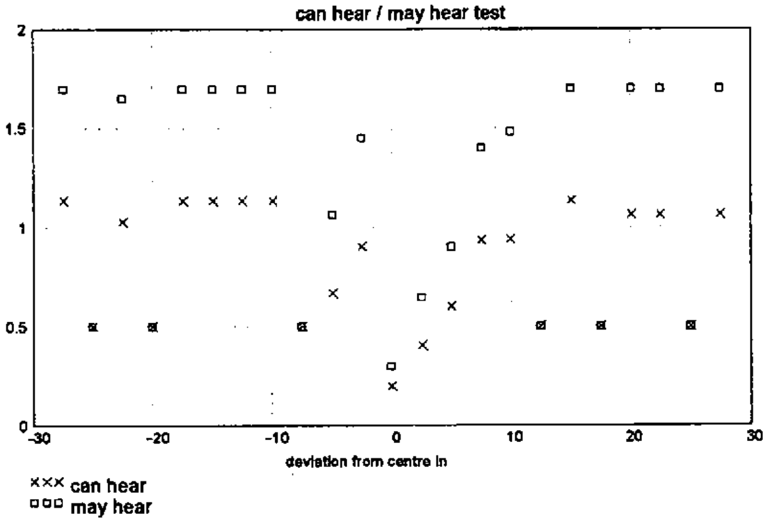


Figure 2. Significance of results being different to p_r=0.5 for Test 2.

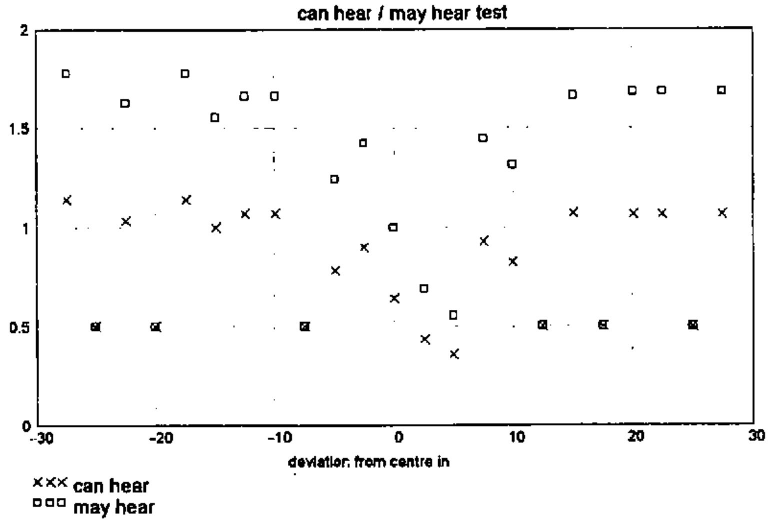


Figure 3. Significance of results being different to p_r=0.5 for Test 3a.

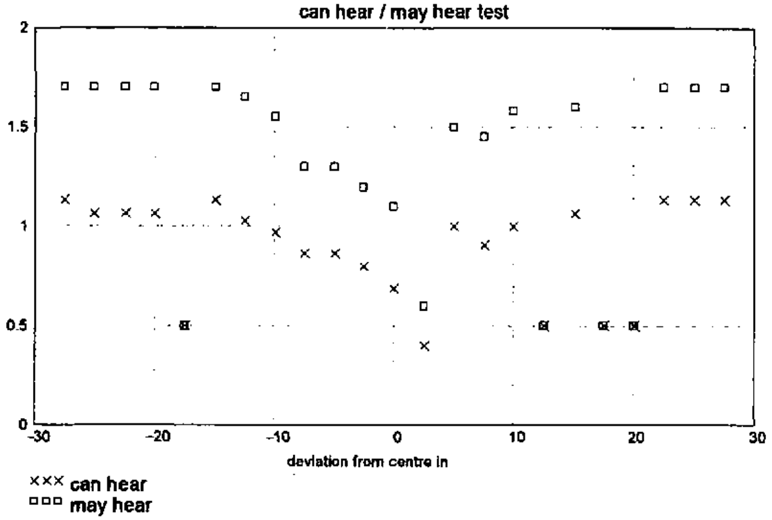


Figure 4. Significance of results being different to p_r=0.5 for Test 4.

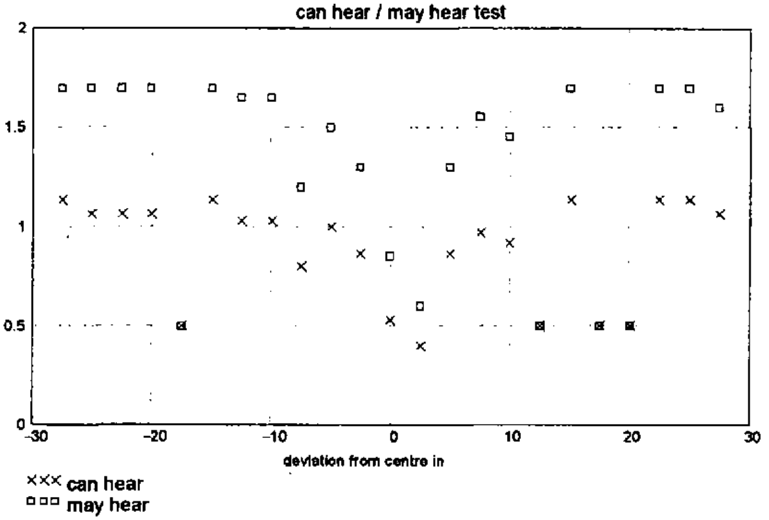


Figure 5. Significance of results being different to p_r=0.5 for Test 5.

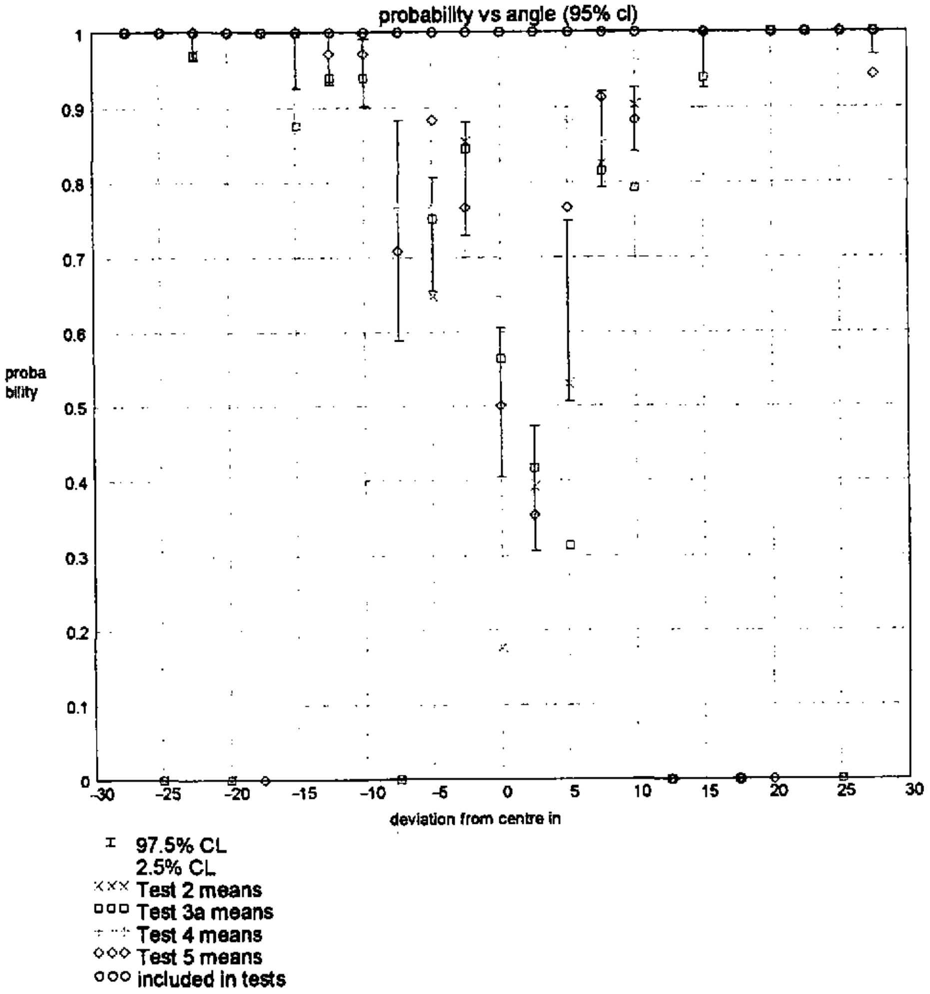


Figure 6. Mean results with 95% confidence limits for tests 2 through 5.

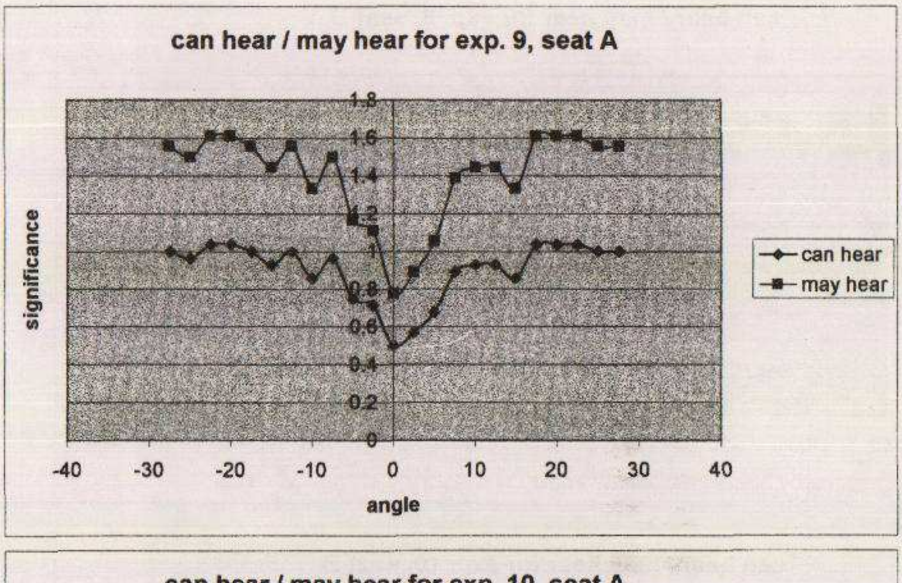


Figure 7. Comparing results from seating position A

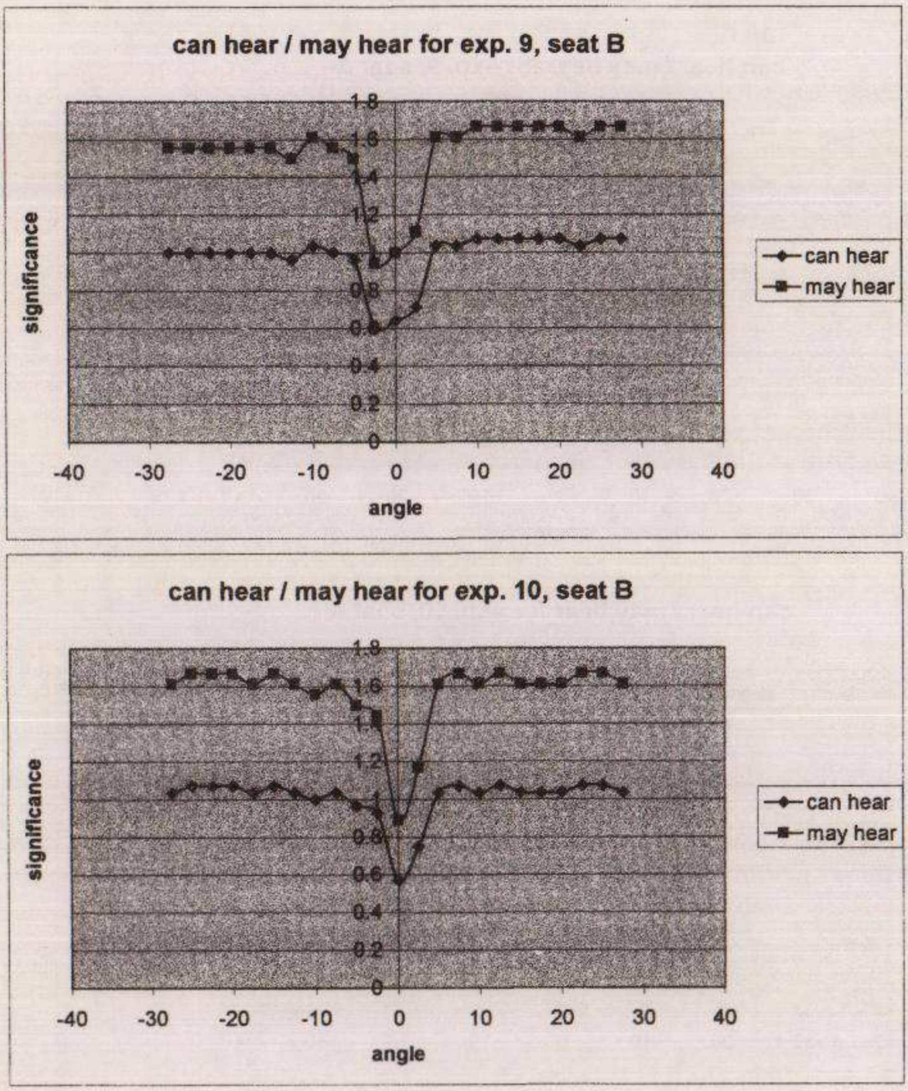
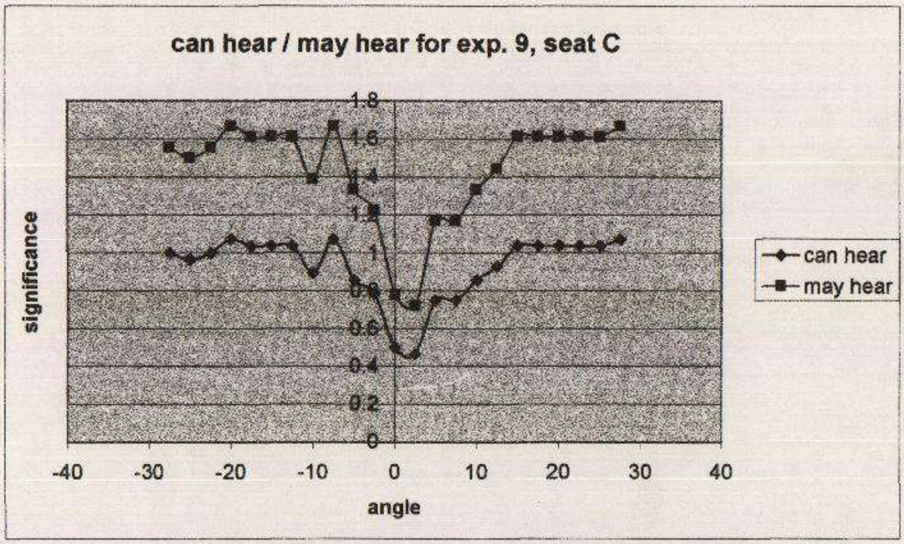


Figure 8. Comparing results from seating position B



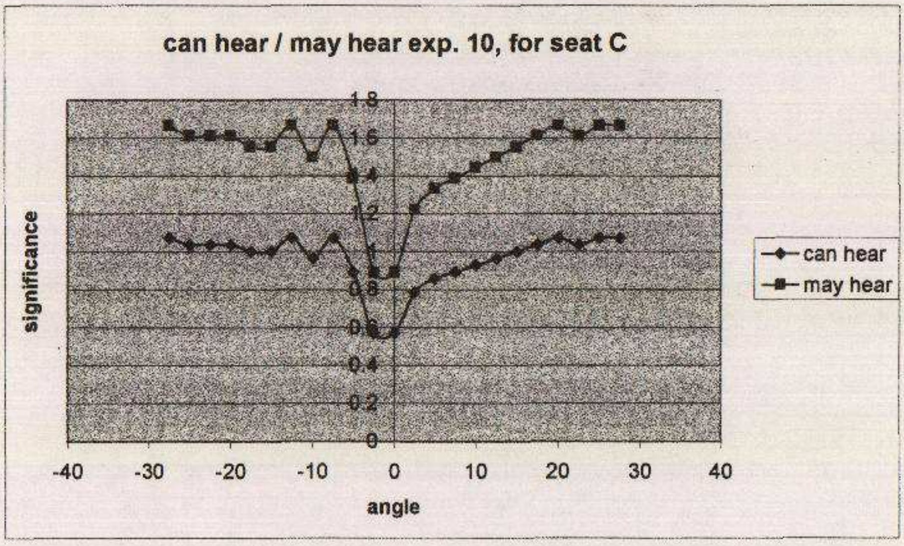
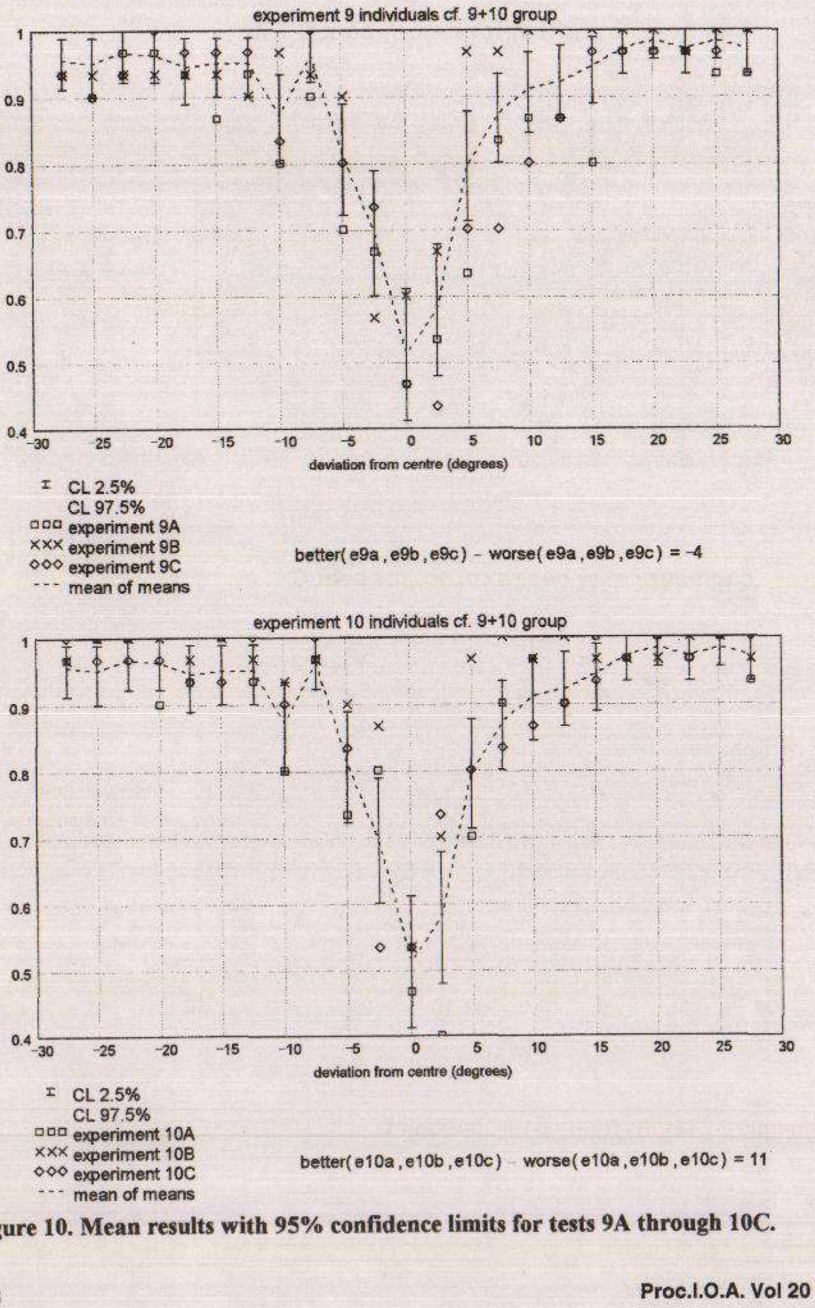


Figure 9. Comparing results from seating position C



better(e9a , e9b , e9c) - worse(e9a , e9b , e9c) = -4

better(e10a , e10b , e10c) - worse(e10a , e10b , e10c) = 11

Figure 10. Mean results with 95% confidence limits for tests 9A through 10C.

experiment	description		two-sided p=0.5							
		97.5%	2.5%	2.5%	97.5%	mean left	mean rignt	difference	mean	rank
2	Tiled DML in wall	-10	0	0	10	-5	5	10	5	2
3a	Mission two-way box	-10	2.5	5	15	-3.75	10	13.75	6.875	4
4	Tiled DML off wall	-10	2.5	2.5	10	-3.75	6.25	10	5	2
5	Tiled DML + woofer hybrid	-10	0	2.5	7.5	-5	5	10	5	2

			can hear / may hear test							
experiment	description	can	may	may	can	mean left	mean rignt	difference	mean	rank
2	Tiled DML in wall	-7.5	-2.5	5	10	-5	7.5	12.5	6.25	2
3a	Mission two-way box	-10	0	7.5	12.5	-5	10	15	7.5	4
4	Tiled DML off wall	-12.5	0	5	5	-6.25	5	11.25	5.625	1
5	Tiled DML + woofer hybrid	-10	-2.5	5	10	-6.25	7.5	13.75	6.875	3

experiment	description		ranking by significance		
		better	worse	delta	rank
2	Tiled DML in wall	0	2	-2	3
3a	Mission two-way box	0	3	-3	4
4	Tiled DML off wall	2	1	1	1.5
5	Tiled DML + woofer hybrid	2	1	1	1.5

Figure 11. **Overall results summary for listening tests 2 through 5**

experiment	description		compare errors to group mean			two-sided significance testing		
		N	total errors	probability	std. Dev.	t statistic	95%	98%
2	Tiled DML in wall	476	85	0.179	0.383	1.91	NOT SIG.	NOT SIG.
3a	Mission two-way box	448	82	0.183	0.387	2.08	WORSE	NOT SIG.
4	Tiled DML off wall	476	50	0.105	0.307	-2.64	BETTER	BETTER
5	Tiled DML + woofer hybrid	476	55	0.116	0.320	-2.01	BETTER	NOT SIG.
	group totals	1876	272	0.145				

Figure 12. **Significance testing for listening tests 2 through 5**

Notes:

The "tiled DML" was demonstrated at the 103rd AES convention in New York
The "tiled DML + woofer hybrid" was demonstrated during the 1998 CES in Las Vegas
The "Mission two-way box" is the 752 model. Mission is a sister company of NXT

	experiment 9A	experiment 9B	experiment 9C	experiment 10A	experiment 10B	experiment 10C
two-sided pr=0.75						
97.50%	-17.5	-7.5	-12.5	-17.5	-7.5	-12.5
2.50%	0	0	0	0	0	-2.5
2.50%	2.5	2.5	2.5	2.5	0	0
97.50%	17.5	5	15	15	5	15
mean width	9.375	3.75	7.5	8.75	3.125	7.5
centre	1.25	1.25	1.25	1.25	0	-1.25
ranking	2	2	1.5	1	1	1.5
can / may hear						
can	-12.5	-7.5	-12.5	-10	-7.5	-10
may	-2.5	-5	-2.5	-2.5	-2.5	-5
may	5	5	5	5	5	2.5
can	17.5	5	15	10	5	15
mean width	9.375	5.625	8.75	6.875	5	8.125
centre	1.25	0	1.25	1.25	1.25	-1.25
ranking	2	2	2	1	1	1

Figure 13. simple ranking by significance tests for the two loudspeaker pairs in tests 9 and 10

compare errors	seat A	seat B	seat C	sum
total errors for 9	110	59	95	264
total errors for 10	83	42	73	198
difference	27	17	22	66
N	690	690	690	2070
μ1	0.159	0.086	0.138	0.128
μ2	0.120	0.061	0.106	0.096
s	12.855	9.657	12.124	20.229
t statistic	2.100	1.760	1.815	3.263
significant at 99%	NO	NO	NO	YES
significant at 97.5%	YES	NO	NO	YES
significant at 95%	YES	YES	YES	YES

(Mission 752 two-way hi-fi speakers[2])

(Cyrus concept NXT hybrid tower[1,2])

Figure 14. Significance testing of differences in total number of errors in tests 9 and 10

Notes:

[1] The "Cyrus NXT hybrid" combines four DMLs with a cone bass unit, and was demonstrated during the 1998 CES in Las Vegas. It is same as the "Tiled DML + woofer hybrid" used in test 5.

[2] Mission, Cyrus and NXT are trademarks of Verity plc, the parent company of NXT.