Enlarging the Sweet-Spot for Stereophony by Time/Intensity Trading

Preprint 3473 (C1-1)

Ronald M.Aarts
Philips Research Laboratories
Eindhoven, Netherlands

Presented at the 94th Convention
1993 March 16–19
Berlin

AN AUDIO ENGINEERING SOCIETY PREPRINT

Enlarging the sweet spot for stereophony by time/intensity trading

Ronald M. Aarts
Philips Research Laboratories
5600 JA, Eindhoven, The Netherlands

The degradation of the stereophonic illusion due to off-center listening is investigated, both for a small and a wide loudspeaker bases. The experimental results for video applications are compared with the case of a much wider loudspeaker base, as in audio-only applications. The optimum loudspeaker radiation patterns are discussed obeying the time/intensity trading concept.

1 Introduction

Since the first demonstration of an electroacoustic reproduction system [1] with the aim of achieving a realistic sound reproduction, increasing efforts have been directed at improving the stereophonic effect [2, 3, 4] and the description of it [5, 6, 7, 8]. Generally, it is considered as a serious artefact of the present stereo system that the listener is aware of the stereophonic illusion only within a limited region. If the head is moved away from the perpendicular on the center of the base of the loudspeakers, then the stereo effect decreases. This off-center listening problem becomes even more serious when the distance between the loudspeakers is not very large in comparison with the deviations from the center position, as in multi-media PC monitor and (HD)TV applications; the latter normally has a wider base but in some cases a smaller base is desired. In the following some experiments are described with the aim of investigating the influence of off-center listening for normal stereo set-ups and for small-base stereo set-ups, as in HDTV and multimedia applications.

Differences between the ear input signals are necessary in order for auditory events to appear laterally, away from the median plane. Interaural time (ITD) and level differences (ILD) are the main parameters for evoking the illusion of a phantom source. A frequently used technique for the measurement of the equivalence between ILD and ITD are trading experiments, that is asking a subject what time difference is equivalent to a certain level difference. A large number of experiments have been reported, see [7] for an overview. However, most are performed by using headphones, or if loudspeakers are used, by applying an artificial delay between them and leaving the subject in the center position. It is the aim of the present experiments to apply natural time differences only, by using a set-up such as occurs in practical situations for normal audio listening and for a (HD)TV sound environment.

2 Experiments

The set-up of the experiment is shown in Figure 1. In an anechoic chamber a pair of two-way (in the horizontal plane) omnidirectional loudspeakers were positioned 2.5 metres apart. These loudspeakers were constructed by mounting the drivers coaxially, with their axes perpendicular to the horizontal plane. Between the two loudspeakers a third loudspeaker was mounted, hereafter referred as the center loudspeaker. The three speakers were mounted at 1.62m, the average ear height of the standing subjects. An arc of seven numbered positions for the subject to stand in was marked in such a way that the subject was always 2.19 metres away from the central loudspeaker. With a DSP and a small computer the set-up, as in Figure 2, was used to control the experiment. The computer drew up a random order of positions for the subject to stand in, in such a way that the experiment was balanced (i.e. all angles had an equal number of occurrences per subject). The difference in volume of the right and left loudspeaker at the start of each run was randomly chosen by the computer in such a way that the total electrical power at the center position was always the same. Throughout the experiment the subject kept his head pointing straight forward (i.e. with his shoulders parallel to the wall). The subjects were given a note about the experiment, and the procedure to be followed was made clear to them before the experiment was started. The task for the subjects was to adjust the level of the right and left loudspeakers until the phantom source was located in the center position. A switching facility was included that sent the signal to the left and right loudspeakers, or to the central loudspeaker, so that at any time during the experiment the subject could switch to the signal coming from the central loudspeaker as a reference of the center position.

The experiment leader indicated to the subject which position to start at. The subject was presented with a box and familiarized with it. Using this box he could control the volume of the left and right loudspeakers, while the total electrical power fed to both loudspeakers was kept constant. The subject could switch to the signal coming from the central loudspeaker only, and back to the left and right loudspeakers, and indicate when he thought the phantom source was at the same position as the center loudspeaker. The subject was given as much time as was required. When the volumes had been adjusted so that the subject felt that the sound was as though it was coming from the central position, he pressed the button marked 'OK' and the difference in level between the right and left loudspeaker was stored in the computer.

This run was then repeated another three times. At the end of the fourth run the experiment leader indicated the new position for the subject and the whole process was repeated until all 7 positions had been taken up.

Table 1: Analysis of Variance Experiment 1.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.59	0.99
angles	27070	6	409	0.0		
subjects	151	9	13	0.6E-14		
interaction	958	54	13	0.4E-42		
error	278	210				
TOTAL	28457	279				

3 Experiment 1

3.1 Conditions

The experiment was carried out using pink noise from a CD. The level at the center position at ear height was about 75dB. The experiment was carried out in an anechoic chamber, with sufficient light for the three loudspeakers to be clearly visible.

3.2 Subjects

Ten subjects with normal hearing, four female and six male, participated in the experiment. The average age of the subjects was 25 (std. dev. 4.4 years). Before the experiments, all the subjects' threshold levels were determined by pure tone audiometry. All the subjects had a hearing loss of less than 10 dB; the threshold differences between the left and right ears of each subject were also less than 10 dB. They were carefully and precisely instructed about their task.

3.3 Results

The level differences between the right (R) and left (L) loudspeakers are plotted in Figure 3. The vertical bars indicate the variance over all 10 subjects with the four replications per subject. In looking at the data two features appear. Firstly, the variance per angle is rather large. This is caused by the large variability among the subjects. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 1.
The entry R^2 shows that 99% of the variance can be explained by the factors angles and subjects and their interaction. The variance of each of the subjects is remarkably low. They can reproduce the task very well, but do not agree with each other. The entry SIG shows the significance level and appears to be high for all factors.
The second feature is the deviation of 2dB at the center position, as can been seen in Figure 3. One would expect this to be 0dB. Initially a psychological reason was expected,

Table 2: Analysis of Variance Experiment 2.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.23	0.99
angles	44878	6	5777	0.00		
subjects	140	9	12	1.0E-13		
interaction	1028	54	14	0.6E-45		
error	272	210				
TOTAL	46318	279				

because the experiment was performed in the corner of an anechoic chamber with the corner to the left of the subject. However, in the third experiment the set-up was made symmetrical, but the difference remained the same. This level difference is also reported in [9, 10, 11]. However the latter considered it to be an experimental artefact. The entry GM in Table 1 is the grand mean of all 280 observations and appears to be very close to 0 dB, as one might expect for equal deviations to the right and left of the median plane, indicating that there is a good symmetry between the left and right plane, except for the center position itself.
All subjects reported a broadening of the phantom source for increasing angles from the center position. At the extreme positions of the subjects, the space between the loudspeakers was filled by the image. This effect is known as image broadening, see [7, 12].

4 Experiment 2

The same set-up and conditions as in the former experiment were used. The only difference was the stimulus. A female voice from CD [13] was used as the stimulus, the duration of the excerpt was about 20 sec. and was then repeated over and over.

4.1 Results

The results are depicted in Figure 3. The figure shows that the results are very similar to those obtained from the pink noise stimulus. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 2.

5 Experiment 3

The aim of this experiment was to study the time/intensity trading at a very close listening distance and a relatively small loudspeaker base, as would occur in Multi Media PC and

Table 3: Analysis of Variance Experiment 3.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.36	0.996
angles	23392	4	9027	0.0E+00		
subjects	191	8	37	0.2E-29		
interaction	757	32	37	0.7E-51		
error	87	135				
TOTAL	24427	179				

TV applications. The set-up is depicted in Figure 4. The switching facility was the same as described before see Figure 2.

5.1 Conditions

The conditions were adapted to the monitor application. A 16:9-aspect-ratio monitor with a 36" screen was used, without displaying a picture. This was done to make the localization more precise than with a picture, because it is known that when the seen and heard locations of an object are to be discrepant, the sound is perceived near or at the seen location. The experiment was performed in a studio using the same female voice as in the 2nd experiment. The subjects were seated in a chair in five different positions. Each position was repeated four times.

5.2 Subjects

Nine subjects out of the set of the previous experiments performed the test.

5.3 Results

The level differences between the right and left loudspeakers were the outcome of the experiment and are plotted in Figure 5. Again, a large variance can been observed and a bias at the center position of almost 2dB appears. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 3 .

The variance within a run for each subject is for this experiment even less then for the previous experiment, 99.6% of the variance being caused by the factors and their interaction. The subjects are consistent, although there is not a broad consensus among the subjects.

6 Experiment 4

The aim of this experiment was to study the trading for three reference positions, so the subject's task was to direct the auditory event in three directions, $+30°$, $0°$ and $-30°$. The set-up of the experiment is similar to that shown in Figure 1, which was used in the first two experiments. However, the subject position was fixed in front of the center loudspeaker and the delay between the loudspeakers was performed by electronic means.

6.1 Conditions

The subjects were seated in a chair, position 4 in Figure 1, with their head clamped. The task was similar to that in the first three experiments. One of the three loudspeakers (randomly chosen) started radiating pink noise. It was the reference loudspeaker/direction for that run and was switched off after two seconds. The subject's task was to adjust the balance between the outer loudspeakers in such a way that it gave the same direction as the reference one. There were 2 runs per subject per delay value and 4 delay values per angle, giving 24 runs per subject.

6.2 Results

The level differences between the right and left loudspeakers were the outcome of the experiment and are plotted in Figure 6. Although the head was clamped, a very large variance can been observed, especially for the two outer reference positions. Again, a bias at the center position appears.

7 Discussion

The data of the first three experiments are converted to exclude the intensity difference caused by different path lengths from left and right loudspeaker to the head, r_{-1} and r_1 respectively. This is done by adding the factor $20\log(r_{-1}/r_1)$ to the R-L values. According to the time-intensity trading concept, the resulting level differences are caused by time differences only. They are plotted in Figure 7, where $\Delta T = (r_1 - r_{-1})/c$. In order to compare the $\Delta T/\Delta I$values, some other values are also plotted in Figure 7. A convenient way to parametrize this set of curves is

$$\Delta L = b\tanh(a\Delta T) \tag{1}$$

where $a \sim 0.8$ and $b \sim 10$. An enormous range of trading ratios has been reported, ranging from 1.7 μsec/dB [14] to 500 μsec/dB. It is known that many factors affect the trading relation. The ILD and the ITD interact in a complex fashion: Some of these features of this interaction may be predicted on physiological grounds. The nature of the stimulus employed and the level at which it is presented are probably the most influential. The smallest ratios are reported for low-frequency tones at high levels, the largest with pulses at low intensity levels. However, there are also important psychological factors. One of these may be conflicting intersensory information. When spatial information in two modalities is available, the information in one of the modalities is partly or wholly discarded in favour of that in the other. When spatial information in light and sound is made to conflict, vision dominates. This phenomenon has been called the ventriloquism effect [15]. It can be important for sound for television, see for example [16, 17, 18], but may also influence the trading experiments; at large angles the subjects can see only one loudspeaker when they are facing forward. The paradigm of the experiment may affect the results, even if the ILDs and ITDs are the same as in another experiment. This was the reason for using a third loudspeaker in the center, so it could serve both as an acoustic and as a visual reference.

8 Application

The data gathered by the experiments discussed before was used to derive an optimal polar pattern for loudspeakers. A pair of loudspeaker cabinets were equipped with a pair of drivers for each frequency range. The drivers were fed by digital filters in such a way that both loudspeakers together radiated sound in a broad range of listening positions in accordance with the time/intensity trading results. An example of a calculated pattern is plotted in Figure 8. The other loudspeaker has the same pattern, but mirrored on the main axis. Informal listening tests showed a pronounced improvement for the optimized patterns in comparison with ordinary loudspeakers.

9 Conclusions

The results for pink noise and the female voice as stimuli are very similar. Time/intensity trading beyond about 1ms is not possible, the phantom source then broadens as the time difference increases. The results exhibit a large variability between the subjects, but a low variance within the subjects. The trading ratios for (HD)TV applications at a close viewing distance and a small loudspeaker base are much steeper than for a wider stereo base in the case of normal audio applications.

References

[1] The telephone at the Paris Opera. Scientific American, December 1881.

[2] A.D. Blumlein. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems. *J. Audio Eng. Soc.*, 6(2):91–99, 1958. British Patent Specification 394,325 (1931).

[3] B.B. Bauer. Broadening the area of stereophonic perception. *J. Audio Eng. Soc.*, 8(2):91–94, April 1960.

[4] M.F. Davis. Loudspeaker systems with optimized wide-listening-area imaging. *J. Audio Eng. Soc.*, 35(11):888–896, 1987.

[5] K. de Boer. *Stereofonische geluidsweergave (Stereophonic sound reproduction).* PhD thesis, Technical University Delft, December 1940.

[6] N.V. Franssen. *Some considerations on the mechanism of directional hearing.* PhD thesis, Technical University Delft, July 1960.

[7] J. Blauert. *Spatial hearing: The Psychophysics of Human Sound Localizaton.* The MIT Press, 1983.

[8] J. Eargle, editor. *Stereophonic techniques: An anthology of reprinted articles on stereophonic techniques.* Audio Eng. Soc., 1986. ISBN 0-937803-08-1.

[9] A.R. McKenzie and I.H. Flindell. Time/intensity trading in stereophony. In *Proceedings I.O.A.*, pages 117–133. Institute of Acoustics (UK), 1989. Volume 11, part 7.

[10] T.T. Sandel et al. localization of sound from single and paired sources. *J. Acoust. Soc. Am.*, 27(5):842–852, September 1955.

[11] K. Wendt. *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie.* PhD thesis, Technische Hochshule Aachen, February 1963.

[12] M.B. Gardner. Image fusion, broadening, and displacement in sound location. *J. Acoust. Soc. Am.*, 46(2 (part 2)):339–349, 1969.

[13] Sound Quality Assessment Material, (recordings for subjective tests). European Broadcasting Union, 1988. no. 422 204-2, track 49.

[14] J.H. Shaxby and F.H. Gage. Studies in the localization of sound. M.R.C. Spec. Rep. Ser. No.166, 1932.

[15] I.P. Howard and W.B. Templeton. *Human spatial orientation.* J. Wiley, 1966.

[16] T. Hatada, H. Sakata, and H. Kusaka. Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. *SMPTE Journal*, 89:560–569, August 1980.

[17] S. Komiyama. Subjective evaluation of angular displacement between picture and sound directions for HDTV sound systems. *J. Audio Eng. Soc.*, 37(4):210–214, April 1989.

[18] R.B. Welch and D.H. Warren. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3):638–667, 1980.

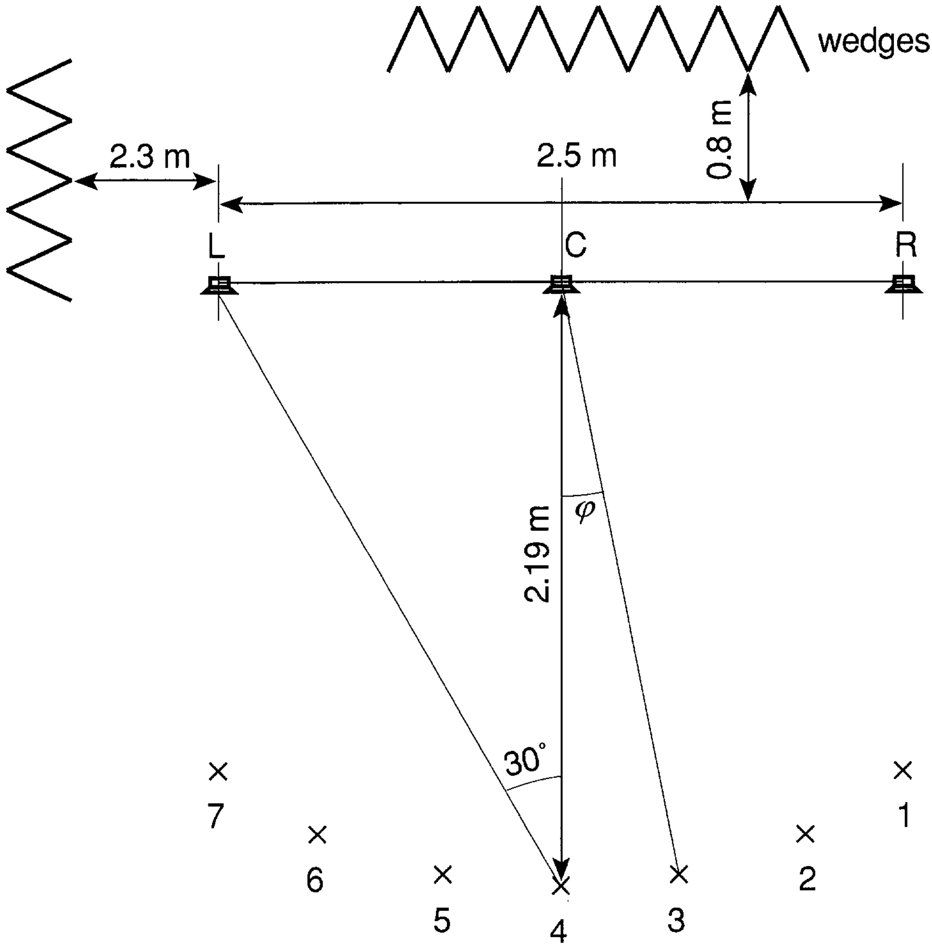

Figure 1: Set-up of the experiment in the anechoic chamber.

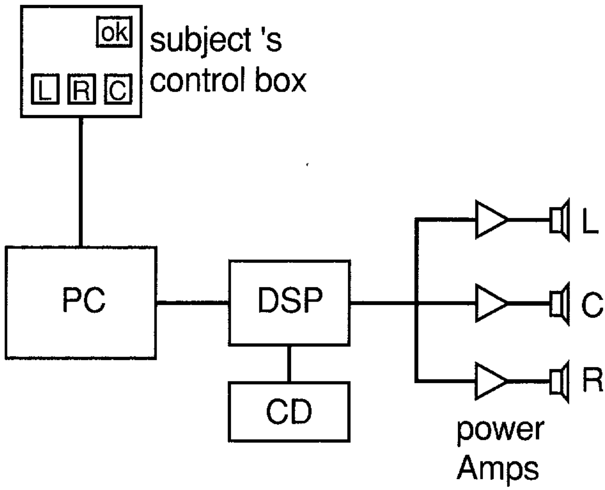

Figure 2: Apparatus for controlling the experiments.

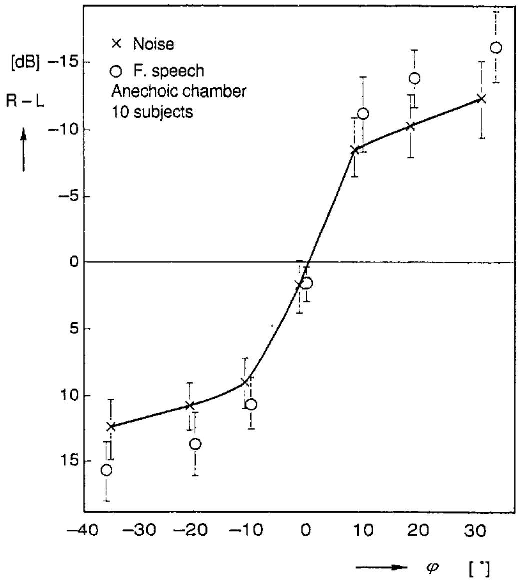

Figure 3: Required level difference between right and left loudspeakers, as a function of the subjects' positions, in order to perceive the phantom image in the center.

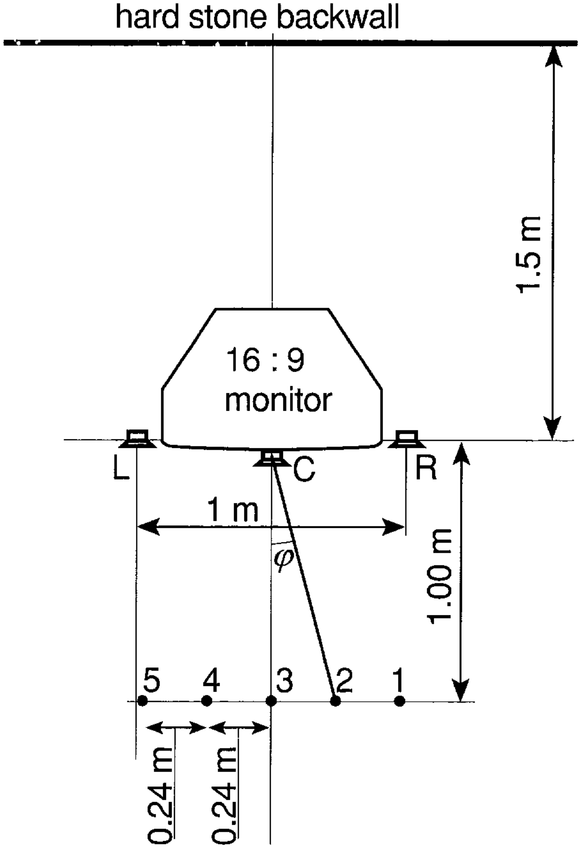

Figure 4: Set-up of the experiment in the studio.

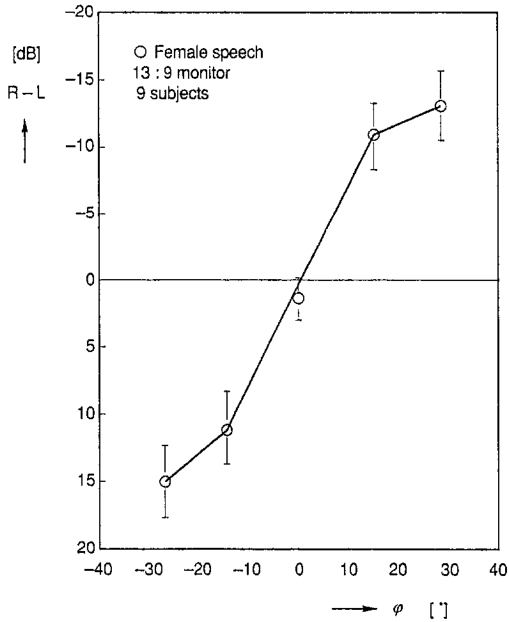

Figure 5: Required level difference between right and left loudspeakers, as a function of the subjects'positions,in order to perceive the phantom image in the center of a TV monitor.

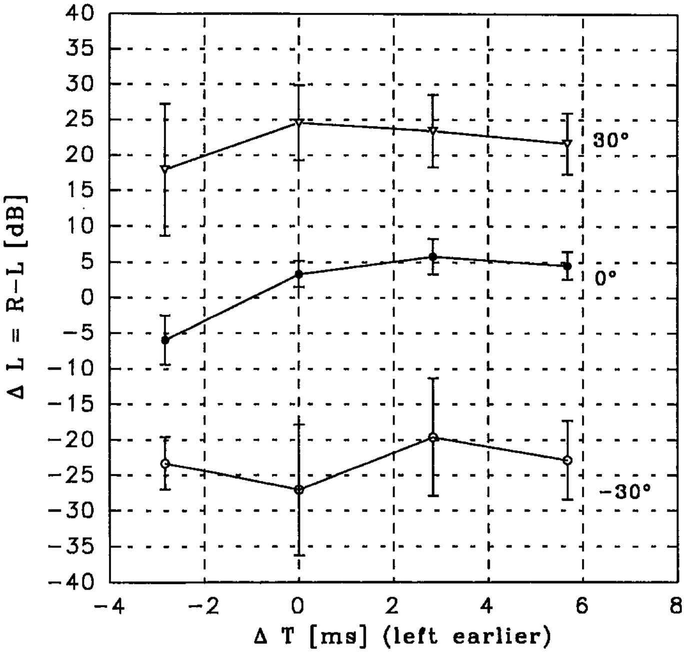

Figure 6: Required level difference between right and left loudspeakers, as a function of the time difference between the loudspeakers, in order to perceive the phantom image in the direction of +30°, 0° or −30°.

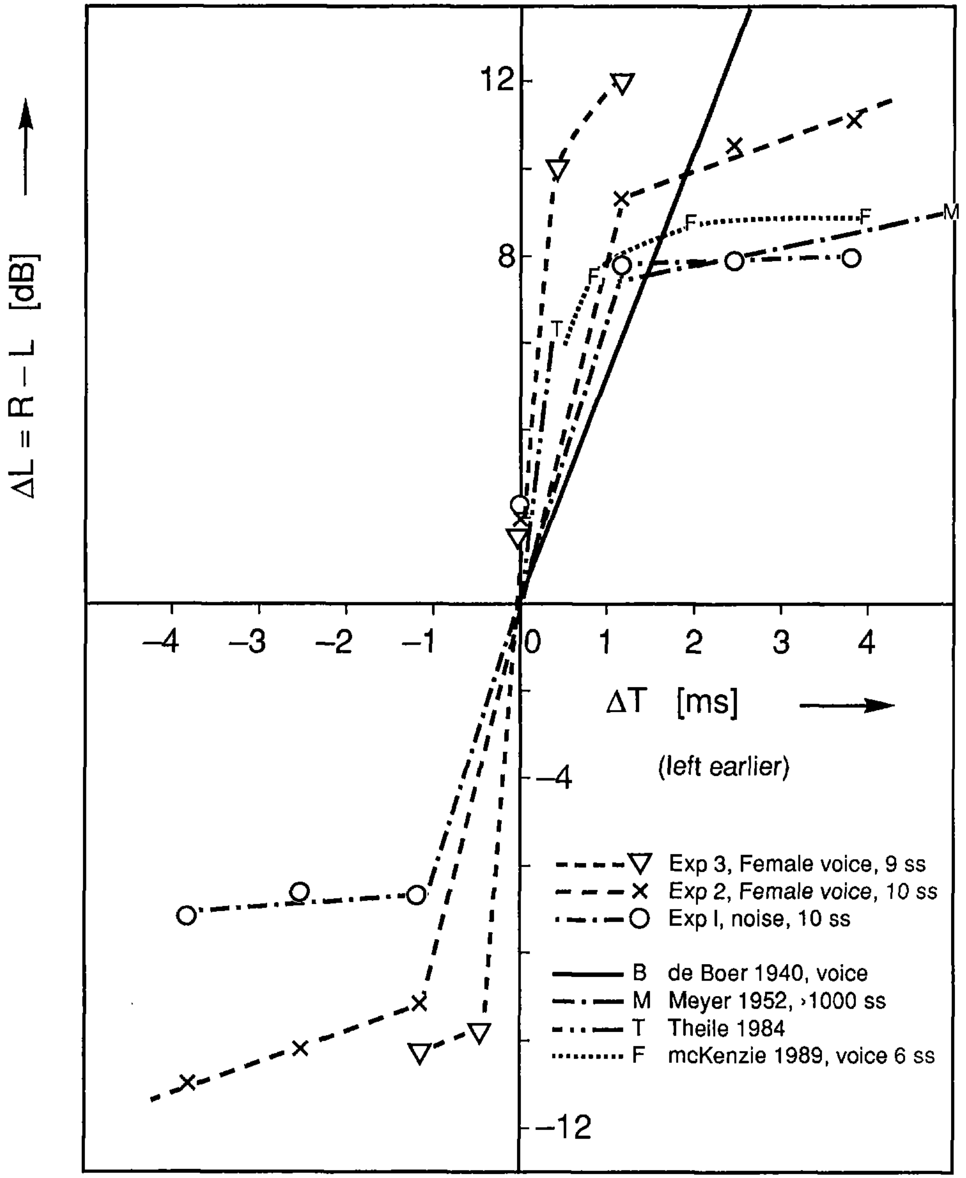

Figure 7: Time/intensity trading results of the experiments in comparison with other experiments from the literature.

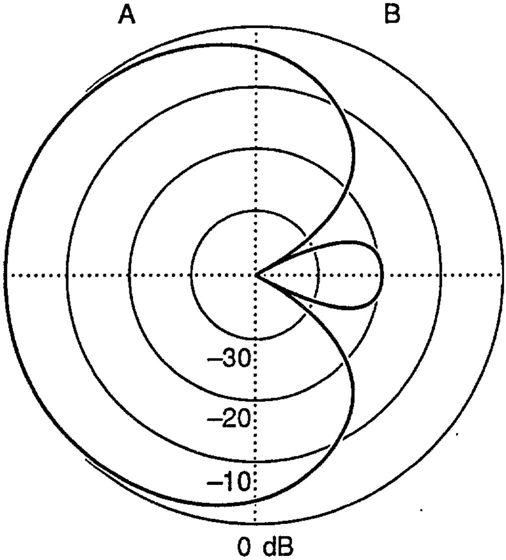

Figure 8: Polar pattern of the optimized loudspeaker.