Multichannel level alignment, part III: The effects of loudspeaker directivity and reproduction bandwidth

Søren Bech[1]
Bang and Olufsen, Struer, Denmark
sbe@bang-olufsen.dk

Nick Zacharov
Nokia Research Center, Speech and Audio Systems Laboratory, Tampere, Finland
Nick.Zacharov@research.nokia.com

Abstract

The correct level alignment of multichannel reproduction systems is critical to the quality of the reproduction. This, the third paper discusses the results of three experiments employing the nine test signals discussed in paper I. Experiments employing six subjects using the method of adjustment have been conducted in a standard listening rooms using a symmetrical loudspeaker set-up. The influence of directivity characteristics of the loudspeakers used for the front and surround channels has been investigated and the influence of high-pass filtering the centre and surround channels has also been examined. The results show that there is no significant influence of directivity characteristics for the front channels. This also applies to the surround channels when using standard loudspeakers that produce both a direct and diffuse sound field component at the listening position. A loudspeaker that only produces a diffuse component will have a significantly different calibration of the surround channel. It is suggested that the effect can be explained by simple changes in the SPL at the listening position. The use of high-pass filtered centre or surround channels does not have a significant influence on the calibration.

1 Introduction

This work forms part of the studies of the Eureka 1653 Medusa (Multichannel Enhancement of Domestic User Stereo Applications) project. The Medusa project is a 3.5 years joint research project with the following partners: British Broadcasting Corporation, The Music Department of the University of Surrey, Nokia Research Centre, Genelec Oy, and Bang & Olufsen A/S.

The purpose of the project is to examine the variables of the domestic multichannel sound system, with and without picture, to carry out the essential optimisation leading to consumer end products. These products will combine the requirements of multichannel reproduction together with the less complex modes of reproduction, such as stereo and mono. These, of necessity, will involve linked studies of programme production and perceptual elements, leading to a single optimised approach to domestic reproduction.

This is the third paper in a series dealing with level calibration of five channel sound systems. The first [1] discussed the generation of a group of experimental calibration signals and the second [2] examined the influence of the signals and position of the loudspeakers.

[1] To whom correspondence should be sent.

The purpose of the relative level calibration is to ensure that the spatial properties of the programme material are reproduced, as closely as possible, in accordance with the intentions of the programme maker. This is well known for two-channel systems intended for reproduction of stereophonic signals. The main requirements for such systems is 1) that the reproduction level at the listening position is identical for both loudspeakers for the same applied signal and 2) that the listener is positioned on the line of symmetry such that the signals from the two loudspeakers arrive at the listening position simultaneously. This will ensure the optimum reproduction of the spatial properties of the programme material.

The purpose of the experiments that are reported in this paper has been to investigate the influence of the directivity characteristics of the individual loudspeakers, and high-pass filtering of the front and surround channel loudspeakers.

2 Experimental strategy

The results reported by Bech [3] suggested that the calibration of a symmetrical multi-channel set-up (identical loudspeakers positioned symmetrically in the room) was identical for the front left and right and surround left and right channels. This was confirmed in a control experiment and reported in Zacharov et al. [2]. The results reported in [2] further showed that the calibration of identical loudspeakers, in the same position, could be repeated between experiments without any significant differences.

The above conclusions were utilised for the present experimental design and it was further decided to divide the experiment into three parts to allow the conclusions of the first experiments to guide the selection of experimental variables and planning of subsequent experiments.

The experimental plan is shown in Table 1. The front left-hand channel is used as the reference for comparisons between experiments. The detailed hypotheses, assuming that experiments can be compared, are shown in Table 2.

3 Physical conditions

3.1 Listening room

The standard listening room of Bang and Olufsen (B&O) was used for these experiments. The details of the room and the set-up are illustrated in table 3 and Figure 2. The room is designed to meet the IEC 268-13 recommendation [4]. The walls are brick and both the floor and ceiling are wooden. The ceiling has been designed as a slit resonator to act as a part of the low-frequency absorption area. The floor is partially carpeted and the room is furnished with a number of chairs, bookshelves and tables, as might be found in a typical domestic environment. The reverberation as a function of one-third octaves is shown in Figure 3.

3.2 Loudspeakers & listening positions

The Quad ESL loudspeaker was chosen because its dipole directivity characteristic allows for a comparison of two extreme situations when used in the surround channel: one when the main lobe is facing the listening position and the second when the main lobe is offset by 90 degrees. The first will result in a sound field at the listening position consisting of a direct and diffuse sound field component, however the second situation only includes a reverberant sound field component. The Genelec 1030A and the BL6000 loudspeakers have more conventional directivity characteristics, however, for mid-frequencies (2 kHz – 12 kHz) the Genelec has a higher directivity index than the average domestic loudspeaker. The directivity indices are shown for all loudspeakers in Figure 1.

The Genelec 1030A loudspeakers were placed at a height of 105 cm, measured from the floor to the centre of the loudspeaker. This was considered to coincide with the mean ear height of listeners. The Quad ESL63 loudspeaker was positioned on a stand specifically made for the Quad ESL63. This raised the geometrical centre of the loudspeaker to a height of 71.8 cm above the floor level and tilted it approximately 10 degrees backwards. The BL6000 loudspeakers are floor-standing loudspeakers with the centre of the tweeter 77.5 cm above floor level.

The listening position was situated at the geometrical centre of the room and the loudspeaker-listener distance was specified as two meters at 0°, ±30°, ±110°, as illustrated in figure 2. In practice this ensured that all speakers are approximately 1 meter from any back wall. This set-up is idealised and in accordance with ITU-R BS 775-1 [5]. For all experiments the loudspeakers were rendered invisible to the subject by means of a visually opaque, but acoustically transparent curtain.

The high-pass filtering (-3 dB @ 150 Hz, 6 dB/octave) of the centre and surround channels was implemented in the electrical set-up as shown in Figure 4. The corner frequency and filter slopes were chosen to represent typical solutions in today’s commercially available systems.

3.3 Signals

A total of nine signals were tested, the development of which has previously been discussed in [1]. The characteristics of these signals are presented in table 4. Signals 1 – 9 were used in experiments 1 and 2 and signals 4, 6, 8, and 9 were used in experiment 3.

The signals were digitally stored on a hard disc based recording system and played back via a 20 bit digital-to-analogue converter at a sampling rate of 44.1 kHz with low-pass filtering at 20 kHz (see also Figure 4).

3.4 Level setting of reference channel and correction for sensitivity of loudspeakers

The reference, centre channel was adjusted to have an equal loudness for all signals in accordance to the Zwicker diffuse field method, as specified within ISO 532 [6]. A level of 20 Sones was found (see [2]) both to provide a comfortable listening condition and allow for sufficient headroom (> 6 dB) in the reproduction system to avoid clipping (see table 5). Note that the calibration of the centre channel was adjusted in experiment 3 where the centre channel is HP filtered. The loudness of the centre channel was thus constant and equal for all signals in all three experiments.

To achieve this a Brüel & Kjær (B&K) 4134 pressure microphone was placed, upward facing, at the centre of the listening position at ear height (1.05m), without the listener's presence. This was connected via a B&K 2609 microphone amplifier to a B&K 2144 analyser with the 7638 loudness module.

The differences in sensitivity between the employed loudspeakers were corrected such that all loudspeakers had the same free-field sensitivity. This was done in accordance with IEC 268 - 5 [7] in the following way: the measurement is made in the anechoic room with a distance of 1.00 m between the microphone and loudspeaker on the axis of the loudspeaker. The noise signal used is band limited pink noise with -3 dB corner frequencies at 200 Hz and 2000 Hz and roll-off slopes of 24 dB/octave.

4 Experimental procedure and evaluation

4.1 Listeners

Six expert listeners from the B&O permanent listening team were employed for these tests. All listeners have long experience in these types of subjective tests, have normal hearing (less than 15 dB deviation for normal) and are trained on a weekly basis in addition to their participation in listening tests.

4.2 Test duration

Each experiment was completed within a period of one week. The tests included 6 blocks for experiment 1 and 4 and 2 blocks for experiments 2 and 3. Each block lasted approximately 20 - 30 minutes. To minimise fatigue on the listeners, a maximum of 2-3 blocks were tested per day.

4.3 Test procedure

A single subject participated in each session and was asked to adjust the level of an individual channel to be subjectively equal to that of the centre channel. A method of adjustment paradigm was employed where the subject was free to switch between the centre channel and the channel to be adjusted. The set-up allowed the subject to make the adjustments in fixed steps of approximately 0.25 dB. The initial level of the channel was randomly set in the range of +2 to +6 dB or -2 to -6 dB. This ensured that the initial level difference between the centre and variable channel was clearly perceivable. The subject was instructed to be facing forward during the entire session.

4.4 Familiarisation and training experiment

Listeners were put through a three-stage procedure for the tests, consisting of

- a familiarisation and training session
- a training experiment
- the main experiments

At the initial stage, listeners were provided with oral and written instructions (see appendix). They were presented with all the signals under consideration, and allowed to consider the task in hand. The use of the switching system was illustrated and listeners were allowed to test the system.

The training experiment is intended to have two functions: 1) train the listener for the task and further familiarise them with the test system and procedure, 2) test for listener reliability. This experiment consisted of a subset of the main experiment with all test signals employed with only the right hand front and surround channels. This data is not employed in the analysis presented in this paper.

4.5 Statistical experimental design

Experiment 1 was designed as full factorial divided into six blocks. The main effects were Signal (degrees of freedom (df) = 8), Channel (df = 4), Subject (df = 5), and Block (df = 5) and the block structure was produced using third-order interactions as generators. The experiment included one repetition so a total of 90 (9*5*2) stimuli were presented to each subject. The Analysis of Variance (ANOVA) model included the main factors, all two-way interactions except Channel*Block & Signal*Block, and the random error.

Experiment 2 had the main effects Signal (df = 8), Channel (df = 2), Subject (df = 5), and Block (df = 3) and the same experimental design strategy as employed for experiment 1.

Experiment 3 had the main effects Signal (df = 3), Channel (df = 2), Subject (df = 5), and Block (df = 1) and the same experimental design strategy as employed for experiment 1.

The initial level of the channel to be adjusted was, as mentioned above, randomly set in a range of +2 to +6 or –2 to –6 dB relative to the centre channel, to provide a clearly perceivable difference between the two channels and to avoid any bias effects. To confirm that the initial level was a random variable it was included as a co-variate in the ANOVA models and tested for significance.

4.6 Objective measurements

In order to facilitate the analysis of the subjective data, a set of physical measurements were performed that would allow for the detailed study of objective metrics of the sound field. To ensure that these measurements were as generic as possible, impulse responses (IR) were collected for each of the reproduction channels at 9 points in the horizontal plane around the listening position, as illustrated in Figure 8 in [2]. The microphone spacing was chosen as 18 cm, as this corresponds to the average separation between the ears. In addition, the central microphone position, in the middle of the listening position was also employed, as during calibration. Having collected IR's, it is possible to calculate a broad range of objective metrics by convolving the IR with the original test signals and applying calibration data. This method was considered more convenient and flexible than making all measurements individually.

5 Results and discussion

5.1 Comparison of front-left channel results between experiments

The experimental strategy is based on the assumption that comparisons can be made between experiments conducted in the same room. The left-front channel has remained the same throughout all experiments to aid this comparison. The results for the left-front channels are shown as a function of signal for experiments 1 and 2 and from [2], in Figure 5. The results shows that there are no significant differences between the level calibrations of the left-front channel for any signal except number 6, where the calibrations in experiments 1&2 are equal, but different from that in [2]. The assumption is therefore confirmed and results can be compared between experiments.

5.2 Experiment 1

The purpose of experiment 1 was to gain an initial insight into the influence of loudspeaker directivity and limitation of the frequency range for the surround channels. The Quad loudspeaker was, as mentioned, chosen to represent directivity characteristics different from the average domestic loudspeaker and secondly to provide the possibility of having only a reverberant field at the listening position.

The data were analysed using the full ANOVA model (see section 4.5) and the residual error variance was checked and found to fulfil[2] the statistical requirements. The non-significant ($p > 5\%$) factors were Block and the interactions between Person*Signal and Person*Block. The results of an ANOVA, using the reduced model is shown in Table 6. The analysis has been conducted using the

[2] It is noted that 9 outliers out of 540 observations had to be excluded to fulfil the requirement of normally distributed residuals. An outlier is defined as an observation that is more than 1.5 times the upper or lower quartile, from the mean value. The exclusion of outliers does only marginally influence the results of the ANOVA and the conclusions are unchanged. The results shown in table 6 are based on the complete data set.

Subject factor as both fixed and random. The same factors were significant in both situations and the results presented in Table 6 are based on a fixed Subject factor. Note that the initial level (Pres_lev in Table 6) is significant and the reported mean values are therefore corrected for this influence. The influence of the initial level is that the calibration increases for increasing initial level.

The results shown in Table 6 mean that the data can be fully represented by plotting the Channel*Signal interaction factor and this is shown in Figure 6. Note that the calibration level (LDIFF) is the electrical signal level of the adjusted channel, relative to the level of the front channel and expressed in dB.

The influence of directivity characteristics in the front channels can be estimated by comparing the left- and right-front channel results in Figure 6 and the right-front and left-front channels in experiment 1 and [2], as shown in Figure 7. The results in Figures 6 and 7 shows that only for signal 6 are there (just) a significant ($p < 5\%$) difference between the calibrations. This indicates that the differences between the Genelec and the Quad directivity characteristics have little influence on the calibration of the front channels for the tested signals.

The influence of band limiting the surround channels can be estimated by comparing the right-surround and the right-surround_2 results in Figure 6. The indication is that there is no significant ($p < 5\%$) influence when using a loudspeaker with directivity characteristics as the Quad. The influence can also be estimated by comparing the results for the left-surround in [2] and the left-surround in the present experiment. These results are shown in Figure 8 and it is seen that there is no significant ($p < 5\%$) difference for any of the signals for a loudspeaker with directivity characteristics as the Genelec. The results thus indicate that the influence of band limiting the surround channels (high-pass) has no significant influence on the level calibration. This is supporting the conclusion by Bech [3], that it is the high frequency part, above 2 kHz of the spectrum that the subject's use for the calibration.

The influence of loudspeaker directivity in the surround channels can be estimated by comparing the left-surround and right-surround_2 results in Figure 6 and the right-surround in experiment 1 and left-surround in [2] as shown in Figure 9. It is seen that for both situations is there a significant ($p < 5\%$) effect for all signals suggesting that the change from the Genelec directivity to the Quad's (in the 90 degree off-set mode) has a strong influence on the level calibration.

The observed influence of the directivity characteristics can be explained as follows: First it is assumed that the total sound pressure level at the listening position has an influence on the subjective calibration level. This is supported by the findings in [2]. Secondly it is assumed that the sound field at the listening position is composed of two parts: the direct sound field from the sound source and the diffuse sound field. The contribution from each part is determined by the interaction between the directivity characteristics of the loudspeaker and the acoustics in the room, as a function of frequency. In the present situation the room factor is constant so the change in calibration level must be attributed to the change in directivity characteristics introduced by the tested loudspeakers.

When the Quad is used in the 90-degree offset mode, the sound field at the listening position only includes the reverberant part of the sound field. This can be compared to a situation where the main lobe of the Quad is facing the listening position. In that case the sound field will include both a direct and a reverberant component. The change in SPL at the listening position when the direct component is removed depends on the ratio of the direct-to-reverberant field. The change in SPL as a function of frequency can be calculated using the formulas below combined with measured values for the reverberation time in the room and the directivity characteristics of the Quad.

$$p_{direct}^2 = \frac{\rho * c * P_a * Q(f)}{4 * \pi * r^2} \qquad [1]$$

$$p^2_{diffuse} = \frac{4 * \rho * c * P_a * T(f)}{0.16 * V} \qquad [2]$$

where p^2 (direct) is the mean squared sound pressure from the direct field component [Pa],
p^2 (diffuse) is the mean squared sound pressure from the diffuse field component [Pa],
c is the speed of sound in m/s,
ρ is the density of air,
P_a is the radiated power,
Q(f) is the directivity factor as function of frequency,
T(f) is the reverberation time as a function of frequency [s],
V is the volume of the room [m^3].

The difference in SPL as a function of frequency for the Quad with the main lobe facing the listening position and the 90-degree offset position is shown in Figure 10.

It is well known [8] that to obtain equal loudness in a free and in a diffuse sound field, respectively requires different sound pressure levels as a function of frequency. Thus to be able to correlate the calculated values to the subjective results, a correction [7] was introduced for the diffuse part of the sound field to compensate for the differences in sensibility. The results based on the corrected values for the diffuse field levels are also shown in Figure 10.

The subjective calibration levels for the Quad in the right-front channel can be compared with those for the right-surround channel as the results in [2] suggested that there is no significant difference in calibration level between the right-front and right-surround channels when using the same loudspeaker type for both channels. These subjective differences are shown in Table 7, together with measured differences in SPL at the listening position for the same signal level feed to the two channels without level calibration applied.

The good correspondence between the measured differences in SPL and the subjective values suggests that the increase in calibration levels for the right surround channel is a simple result of the lower SPL caused by the missing direct sound field component for that channel. The calculated values shown in Figure 8 suggest that it is the frequency range above 2 kHz that is used by the subjects for the calibration. This is in agreement with previous results discussed above and in [2].

To summarise the findings of experiment 1:

- The difference in directivity characteristics between the Quad and the Genelec loudspeakers, when used for the front channels, does not have a significant influence on the subjectively adjusted level calibration for the signals tested,
- The tested HP filtering (3 dB @ 150 Hz, 6 dB/oct.) of the signals to the surround channels does not have a significant influence on the subjectively adjusted level calibration for the signals and loudspeakers tested,
- The difference in directivity characteristics between the Quad in the 90 degrees offset position and the Genelec loudspeakers, when used for the surround channels, does have a significant influence on the subjectively adjusted level calibration for the signals tested. The observed differences in calibration level is probably to compensate for the reduction in the SPL at the listening position, caused by the missing direct sound component of the Quad loudspeaker.

The observed influence of the directivity characteristics for the surround channel is based on the use of the Quad loudspeaker in the 90 degrees offset mode (Quad_90). The loudspeakers have all been adjusted to have the same free-field sensitivity, however it could be argued that the correct calibration for the Quad_90 situation would be based on its power response. However, if the assumption is correct that the total SPL level at the listening position has a strong influence on the subjective level calibration, then it follows that a sensitivity correction based on a power measurement would just change the subjective levels by a fixed amount, determined by the difference between the free-field and power corrections. It can be calculated that the difference between a free-field and power based sensitivity correction is 5.85 dB for the Quad loudspeaker. This means that if the Quad_90 had been sensitivity corrected according to power, the subjective calibrations would have been approximately – 3 dB instead of the observed + 3 dB (see Figure 6 for absolute values).

1.3 Experiment 2

The purpose of experiment 2 was to repeat experiment 1 using a loudspeaker with directivity characteristics similar to the average consumer loudspeaker instead of the Quad loudspeaker. The BL6000 was chosen for that purpose.

It was also decided to do no further tests on the influence of HP filtering in the surround channels. The results of experiment 1 indicated that there was no significant influence of the tested HP filtering of the signals to the surround channels. The result is not likely to be dependent on the type of loudspeaker employed for the surround channel as most loudspeakers will have directivity characteristics in the range covered by the Quad and Genelec loudspeakers (see Figure 1) in the frequency range below 150 Hz.

The data were analysed using the full ANOVA model (see section 4.5) and the residual error variance was checked and found to fulfil the statistical requirements. The non-significant ($p > 5\%$) factors were Channel, Block, and Presentation Level. The results of an ANOVA, using the reduced model is shown in Table 8. The analysis has been conducted using the Subject factor as both fixed and random. The same factors were significant in both situations and the results presented in Table 6 are based on a fixed subject factor.

The results shown in Table 8 mean that the data can be fully represented by plotting the Channel*Signal interaction factor and this is shown in Figure 11.

The influence of directivity characteristics in the front channels can be examined by comparing the results for the right-front channels in experiments 1and 2, shown in Figure 12, and the left- and right-front channels in experiment 2, shown in Figure 11. Both set of results indicate that there is no significant effect of the introduced differences in directivity characteristics.

The influence of directivity characteristics in the surround channels can be examined by comparing the results for the right-surround channel in experiments 1 and 2, shown in Figure 13 and the right-surround channel in experiment 2 and the left-surround in [2], shown in Figure 14. The results shown in Figure 13 confirm the results of experiment 1 as the difference between the Quad in 90 degrees mode in experiment 1 and the BL6000 in experiment 2, produce significant differences in calibration level. The results in Figure 14 further show that the differences between the Genelec and the BL6000 directivity characteristics are not of a magnitude that have a significant influence on the level calibrations.

To summarise the results of experiment 2:

- The difference in directivity characteristics between the Genelec and the BL6000 loudspeakers, and the BL6000 and the Quad when used for the front channels, does not have a significant influence on the subjectively adjusted level calibration for the signals tested,
- The difference in directivity characteristics between the Genelec and the BL6000 loudspeakers when used for the surround channels does not have a significant influence on the subjectively

adjusted level calibration for the signals tested,
- The difference in directivity characteristics between the BL6000 and Quad in the 90 degrees offset mode when used for the surround channels does have a significant influence on the subjectively adjusted level calibration for all the signals tested.

1.4 Experiment 3

The purpose of experiment 3 was to test the influence of HP filtering the centre channel. This will be a common situation in many commercial systems, and especially if the centre loudspeaker is included in an existing television set. It was decided only to test a limited number of signals to limit the experimental effort. Signals 4, 6, 8, and 9 were chosen partly based on the results in experiments 1 and 2 where these signals exhibited the largest sensitivity for the introduced changes, compared to the other signals and partly because they are representative (spectrally) of the whole group of signals.

The data were analysed using the full ANOVA model (see section 4.5) and the residual error variance was tested and found to fulfil the statistical requirements[3]. The significant factors ($p < 5\%$) were Person, Channel and the interaction between Person & Channel.

The influence of the HP filtering can be tested by comparing the calibrations obtained for the same loudspeaker in the same channels with and without a HP filtered centre channel. The results are shown in Figures 15 – 18 and they show that there is no significant effect of introducing a HP filtered centre channel for any of the tested situations.

It is interesting to note that there is, although not significant, a consistent trend that the calibration level decreases when the HP filtered centre is introduced for the BL6000 and Genelec loudspeakers (Figures 15 – 17), and that it increases for the Quad loudspeaker (Figure 18). This suggests a complicated interaction between directivity characteristics and HP filtering of the centre channel. This will be investigated in more detail in the next paper where the results of the objective measurements will be reported. The difference in trend for the two groups of loudspeakers also causes the Channel effect to be a significant factor, as observed in the ANOVA.

1.5 The effects of signal

The ANOVA's for experiments 1 and 2 (tables 6 and 8) showed that there was a significant effect ($p < 5\%$) of signal and for experiment 3 the Signal factor was nearly significant ($p = 8\%$). To examine the signal factor in more detail the results for experiments 1 – 3 have been re-plotted in Figures 19 – 21. The results indicate that signal 6 for the left-front channel in experiments 1 and 2 is the cause of the significant interaction between channel and signal. Signal 6 is also seen to result in a lower calibration level for the left-front in experiment 3. This signal is specifically designed to have constant loudness as a function of frequency (ERB's for this signal). Moore and Glasberg's loudness model [9] have been used to create this signal and it is characterised by a 10 dB boost at low frequencies (< 100 Hz) and 10 – 20 dB for higher frequencies (> 5 kHz). The reader is referred to paper I [1] for specific details. A detailed investigation of the correlation between the subjective calibrations and the characteristics of the signals will be presented in the forthcoming paper IV. It is further noted that none of the commercially available signals (1, 2, 3, 8, and 9) result in significantly different calibration levels for any of the investigated situations.

[3] It is noted that 11 outliers out of 144 observations had to be excluded to fulfil the requirement of normally distributed residuals. An outlier is defined as an observation that is more than 1.5 times the upper or lower quartile, from the mean value. The exclusion of outliers does only marginally influence the results of the ANOVA and the conclusions are unchanged.

6 Summary and conclusions

The purpose of the experiments was to investigate the influence of loudspeaker directivity characteristics in front and surround channels and high-pass filtering of the centre and surround channels, on the subjective level calibration of the individual channels with respect to the centre channel. The experiments were conducted in Bang & Olufsen's listening room using their permanent listening team of six subjects. The loudspeaker and listener arrangement was positioned symmetrical in the room and in accordance with the ITU BS 775-1. Nine signals were examined using three (four) types of directivity characteristics for the individual channels. The influence of high-pass filtering was examined using a high-pass filtered (- 3 dB @ 150 Hz, 6 dB/octave) version of the centre and surround channel. The subjects adjusted the level of the individual channel to be equally loud as the centre channel using a method of adjustment procedure. The loudness level of the centre channel was adjusted to be constant (20 Sones) for all signals in all experiments. The conclusions based on the experimental results and the above conditions are as follows:

Front channels
Three different directivity characteristics according to the Genelec, BL6000, and Quad (in normal mode) loudspeakers were examined and the results have shown that there are no significant differences between the subjective level calibrations of the front channels for the three directivity characteristics.

Surround channels
Three different directivity characteristics according to the Genelec, BL6000, and Quad_90 (the 90 degree offset mode) loudspeakers were examined and the results have shown that there are no significant differences between the subjective level calibrations of the surround channel using the Genelec and BL6000 loudspeakers. There are, however significant differences between the calibrations for the Quad_90 and the two other loudspeakers. The hypothesis has been suggested that these differences are to compensate for differences in SPL at the listening position caused by the differences in directivity characteristics.

The introduced high-pass filtering does not have a significant influence on the subjective level calibrations for the tested loudspeakers and signals.

Centre channel
The introduced high-pass filtering does not have a significant influence on the subjective level calibration of the front or surround channels for the tested signals and loudspeakers.

Influence of signal type
The results have shown that there are no significant differences between the subjective level calibrations as a results of using different signals except when signal six is used for the front channel.

7 Acknowledgements

The authors would like to thank the funding bodies for supporting the Eureka Medusa project, including Tekes (Technology Development Centre of Finland) and Danish DTI. All members of the Medusa project are thanked for their comments and discussion throughout the project so far.

Jan Larsen is thanked for setting up and conducting experiments at B&O and Peter Chapman for performing the objective measurements.

8 References

[1]. Suokuisma P., Zacharov N., Bech S., Multichannel level alignment, part I: Signals and methods, 105th Convention of the Audio Engineering Society, San Francisco, USA, 1998, pre-print 4815.

[2]. Zacharov N., Bech S., Suokuisma P., Multichannel level alignment, part II: The influence of signals and loudspeaker Placement, 105th Convention of the Audio Engineering Society, San Francisco, USA, 1998, pre-print 4816.

[3]. Bech S., Calibration of relative level differences of a domestic multichannel sound reproduction system, J. Audio Eng. Soc., vol. 46, pp 304 – 313, April 1998.

[4]. IEC Recommendation 268-13, "Sound system equipment - Part 13: Listening tests on loudspeakers", International Electrotechnical Commission, Geneva, Switzerland (1985)

[5]. ITU-R Recommendation BS 775-1, Multichannel stereophonic sound system with and without accompanying picture, Geneva, 1994.

[6]. ISO Standard 532, Acoustics - Method for calculating loudness level, International Organisation for Standardisation, 1975

[7]. IEC Recommendation 268-5, "Sound system equipment - Part 5: Loudspeakers", International Electrotechnical Commission, Geneva, Switzerland (1989)

[8]. ISO Standard 454, Acoustics – Relation between sound pressure level of narrow bands of noise in diffuse field and in a frontally incident free field for equal loudness, International Organisation for Standardisation, 1975

[9]. Moore, B. C. J. and Glasberg, B. R., " A revision of Zwicker's Loudness Model", Acustica, 82, pp. 335 – 345, 1996

Figures & Tables

Experiment	Centre channel	Front-left Channel [LF]	Right-front Channel [RF]	Left-surround channel [LSS]	Right-surround channel [RSS]	Right-surround channel-2 [RSS_2]
[2]	1-FBW	1-FBW		1-FBW		
1	1- FBW	1-FBW	2-0-FBW	1-HP	2-90-FBW	2-90-HP
2	1-FBW	1-FBW	3-FBW		3-FBW	
3	1-HP	1-FBW	2-0-FBW		3-FBW	

Table 1 General experimental design. Notations for entries are: 1 is the Genelec 1030A loudspeaker, 2-0 is the Quad ESL63 loudspeaker with the main lobe pointing towards the listening position, 2-90 is the Quad ESL69 loudspeaker with the main lobe angled at 90 degrees towards the listening position, 3 is the BL6000 loudspeaker. FBW means Full Band Width of the signals and HP means High-Pass filtered signal (3 dB @ 150 Hz, 6 dB/octave). The two coloumns for the right surround channel means that the same Quad loudspeaker was used in both situation, but the signal was either FBW or HP filtered. The design used in Zacharov et al. [2] is shown in the first line for comparative reasons.

Experiment	**Channels to be compared**			**Influence to be tested**
1	Right-Front (Quad-FBW)	&	Left-Front in [2] (Genelec-FBW)	Loudspeaker directivity in front channels
1	Left-Front (Genelec-FBW)	&	Right-Front (Quad-FBW)	Loudspeaker directivity in front channels
1	Left-Surround (Genelec-HP)	&	Left-Surround in [2] (Genelec-FBW)	Limitation in frequency range for surround channels
1	Right-Surround (Quad-90-FBW)	&	Right-Surround_2 (Quad-90-HP)	Limitation in frequency range for surround channels
1	Right-Surround (Quad-90-FBW)	&	Left-Surround in [2] (Genelec-FBW)	Loudspeaker directivity in surround channels
1	Left-Surround (Genelec-HP)	&	Right-Surround_2 (Quad-90-HP)	Loudspeaker directivity in surround channels
2	Right-Front (BL6000-FBW)	&	Right-Front in exp. 1 (Quad-FBW)	Loudspeaker directivity in front channels
2	Left-Front (Genelec-FBW)	&	Right-Front (BL6000-FBW)	Loudspeaker directivity in front channels
2	Right-surround (BL6000-FBW)	&	Right-Surround in exp. 1 (Quad-90-FBW)	Loudspeaker directivity in surround channels
2	Right-Surround (BL6000-FBW)	&	Left-Surround in [2] (Genelec-FBW)	Loudspeaker directivity in surround channels
3	Right-Surround (BL6000-FBW)	&	Right-Surround from exp. 2 (BL6000-FBW)	Limitation in frequency range for centre channel
3	Left-Front (Genelec-FBW)	&	Left-Front in exp. 1 (Genelec-FBW)	Limitation in frequency range for centre channel
3	Left-Front (Genelec-FBW)	&	Left-Front in exp. 2 (Genelec-FBW)	Limitation in frequency range for centre channel
3	Right-Front (Quad-FBW)	&	Right-Front from exp. 1 (Quad-FBW)	Limitation in frequency range for centre channel

Table 2 Outline of hypotheses to be tested in experiments 1 – 3.

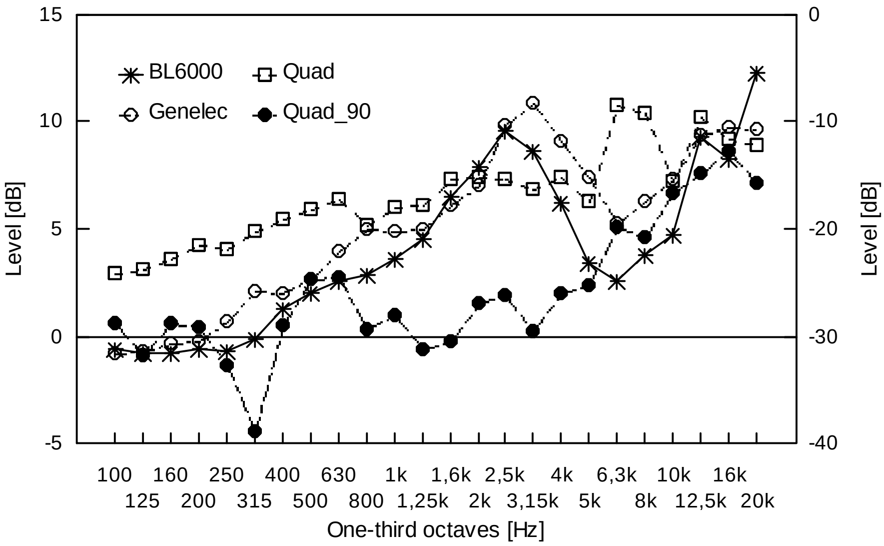

Figure 1 Directivity characteristics for the employed loudspeakers as function of one-third octaves. Note that the right-hand ordinate axis is used for the Quad_90 data. The data are based on free-field measurements of the power and reference axis frequency responses.

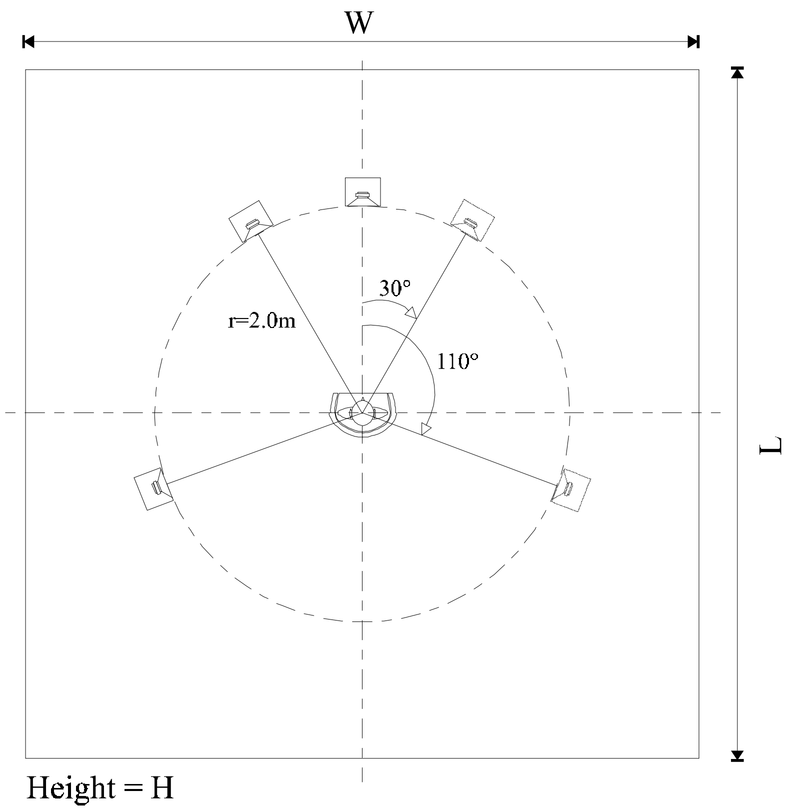

Figure 2. B&O listening room set-up. Loudspeaker icones illustrate the positions of the loudspeakers, and not the actual type employed for the tests. See Table for further details. The listening position was on the line of symmetry in the room.

	B&O Listening room
Volume, (m^3)	82.9
Height, H (m)	2.65
Width, W (m)	5.03
Length, L (m)	6.03
Floor area (m^2)	30.3
Background noise level (dB SPL, A-weighted, fast)	< 35

Table 3. Summary of listening room characteristics

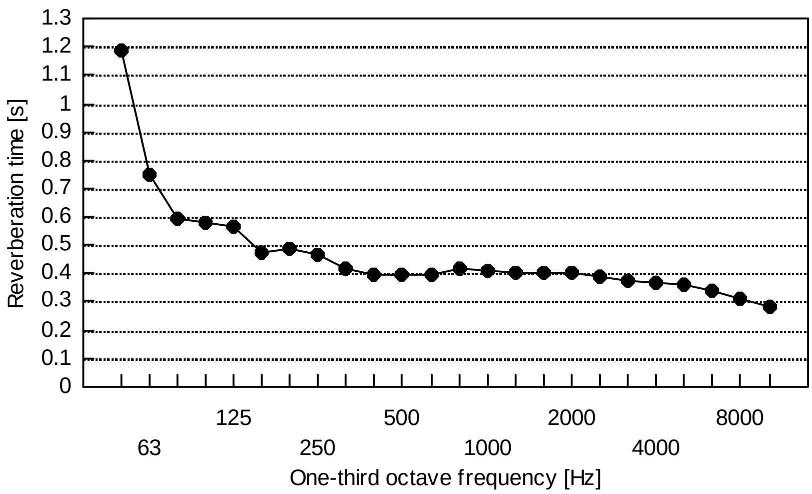

Figure 3 Reverberation time (RT60) as a function of one-third octaves for the B&O listening room.

Signal Name	**High pass filter characteristics**	**Low pass filter characteristics**	**Comments**
	Hz , dB/Oct.	**Hz , dB/Oct.**	
1.	700, 12	700, 6	Commercially available signal
2.	250, 6	500, 6	A signal
3.	500, 18	2k, 18	Commercially available signal
4.			Zwicker constant specific loudness according to ISO 532 (diffuse field)
5.			Zwicker constant specific loudness according to ISO 532 (free field)
6.			Constant specific loudness according to Moore and Glasberg's loudness model [8]
7.			Uniform excitation noise according to Zwicker's loudness model [6]
8.			Pink noise
9.			B-weighted pink noise

Table 4 Characteristics of test signals. See [1] for further details.

Signal name	**Diffuse field loudness**	**SPL's at listening position**	
	Sones (±0.5)	**Linear, Slow**	**A-weighted, slow**
1.	20	67.2	66.5
2.	20	70.0	65.1
3.	20	69.5	69.0
4.	20	70.8	63.5
5.	20	72.9	65.8
6.	20	71.4	*66.5*
7.	20	65.2	*63.8*
8.	20	67.0	62.9
9.	20	65.5	63.6

Table 5 Center channel loudness alignment for experiment 1.

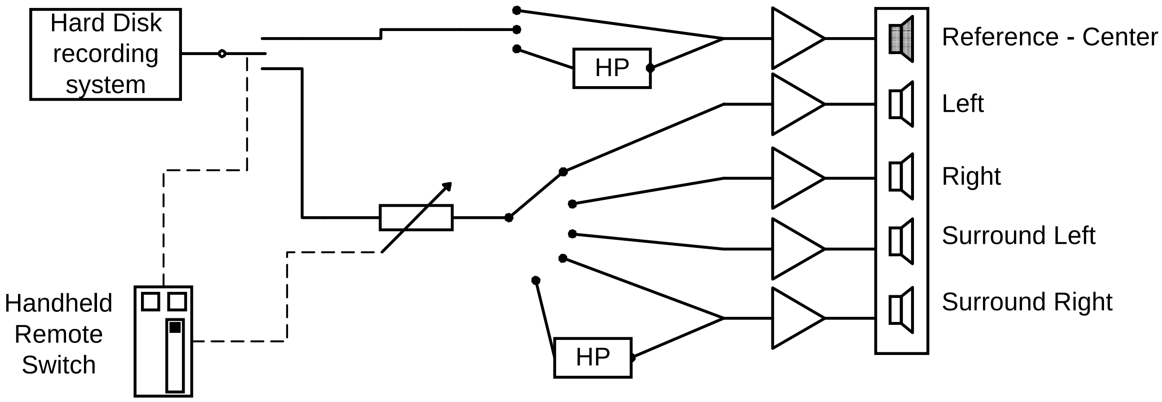

Figure 4 Block diagram of electrical set-up.

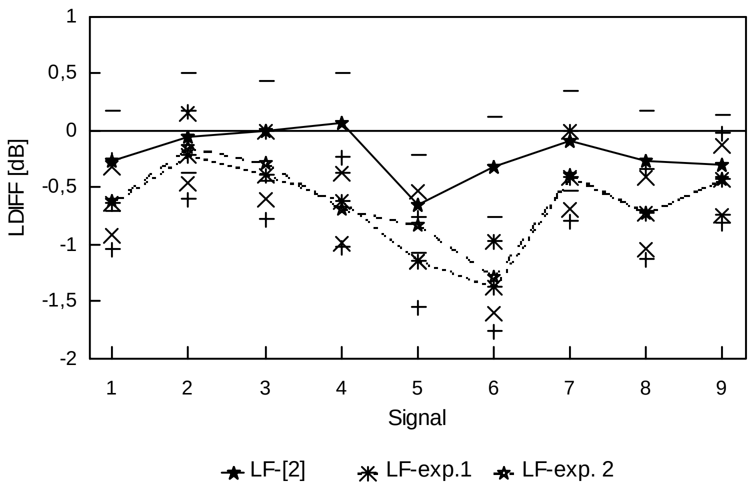

Figure 5 Calibration levels for left-front channel in [2] and in experiments 1 and 2. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

Dependent Variable: LDIFF

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	2420,169[b]	70	34,574	73,075	,000	5115,266	1,000
Intercept	628,742	1	628,742	1328,908	,000	1328,908	1,000
CHANNEL	1949,782	4	487,446	1030,264	,000	4121,058	1,000
PERSON	300,838	5	60,168	127,170	,000	635,851	1,000
SIGNAL	11,120	8	1,390	2,938	,003	23,504	,954
CHANNEL * PERSO	106,808	20	5,340	11,287	,000	225,749	1,000
CHANNEL * SIGNAL	40,338	32	1,261	2,664	,000	85,258	1,000
PRES_LEV	9,222	1	9,222	19,492	,000	19,492	,993
Error	221,896	469	,473				
Total	3267,308	540					
Corrected Total	2642,066	539					

a. Computed using alpha = ,05

b. R Squared = ,916 (Adjusted R Squared = ,903)

Table 6 ANOVA results for experiment 1. The model only includes significant factors. Pres_level is a co-variate representing the initial presentation level (see text for more information).

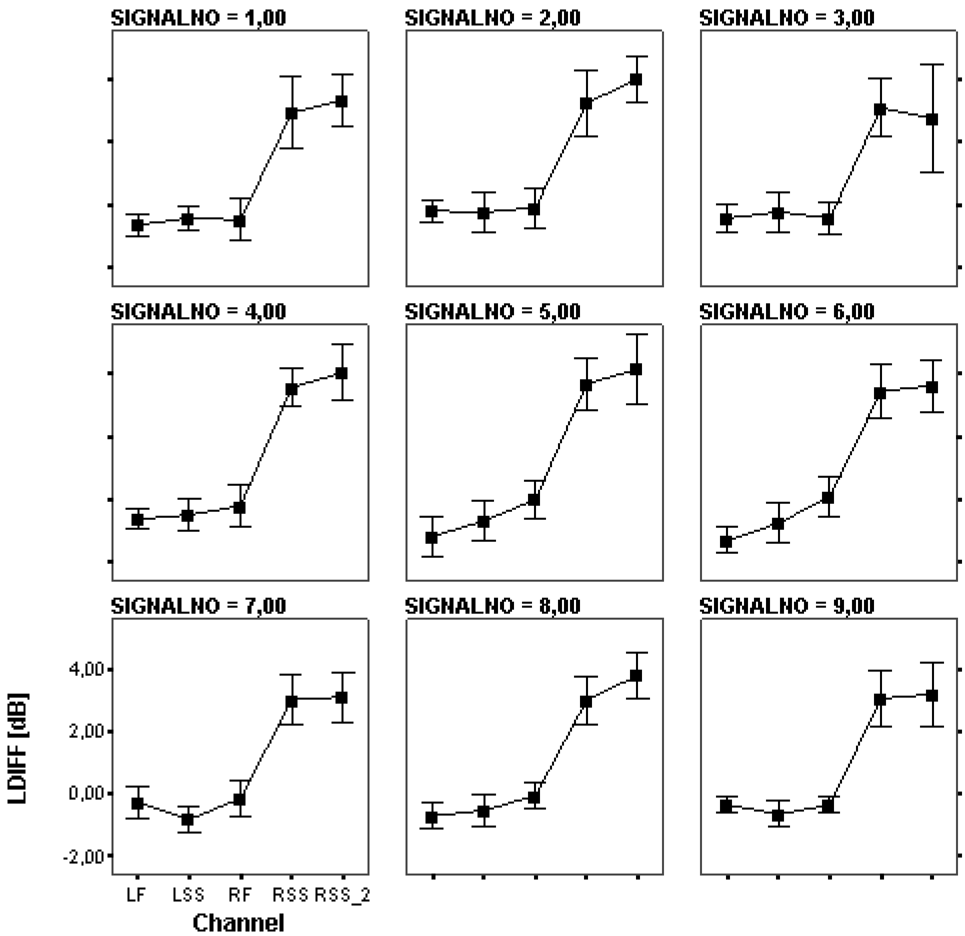

Figure 6 Calibration level for all channels in experiment 1. Shown are mean vaues averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

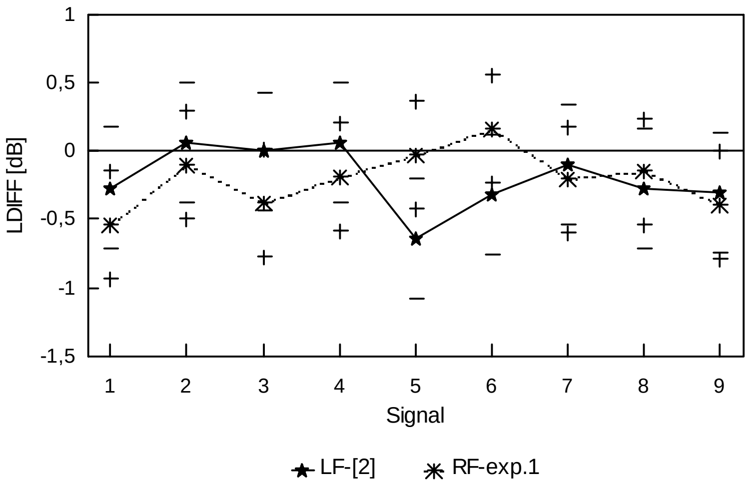

Figure 7 Calibration levels for left-front channel in [2] and right-front channel in experiment 1. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

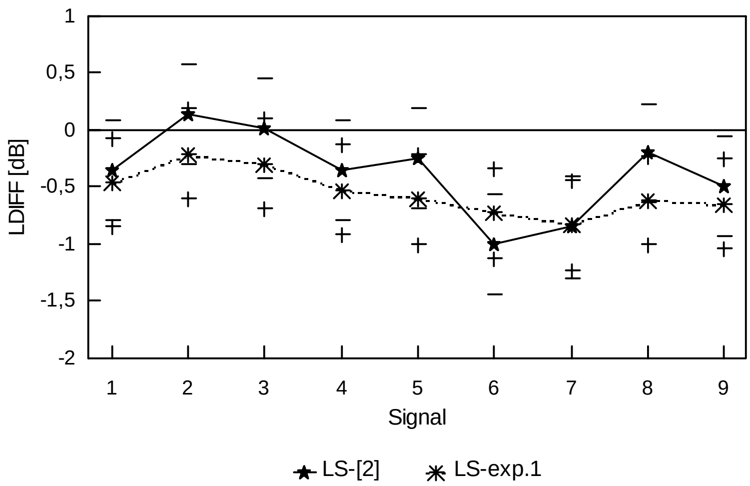

Figure 8 Calibration levels for left-surround channels in experiment 1 and from [2], as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

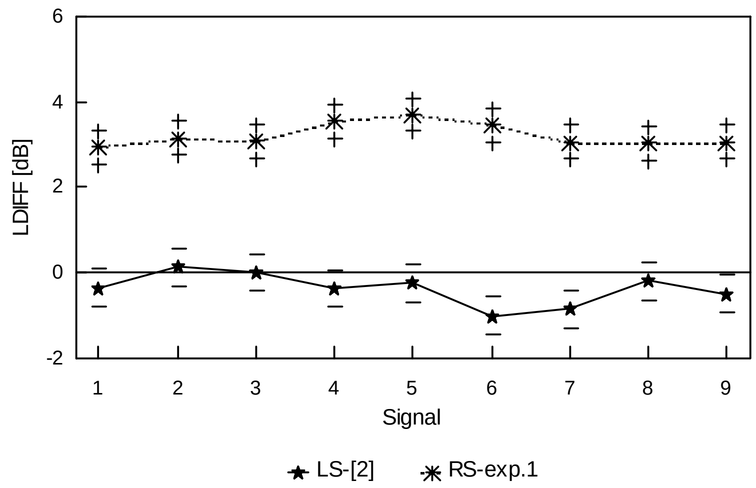

Figure 9 Calibration levels for right-surround in experiment 1 and left-surround in [2], as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

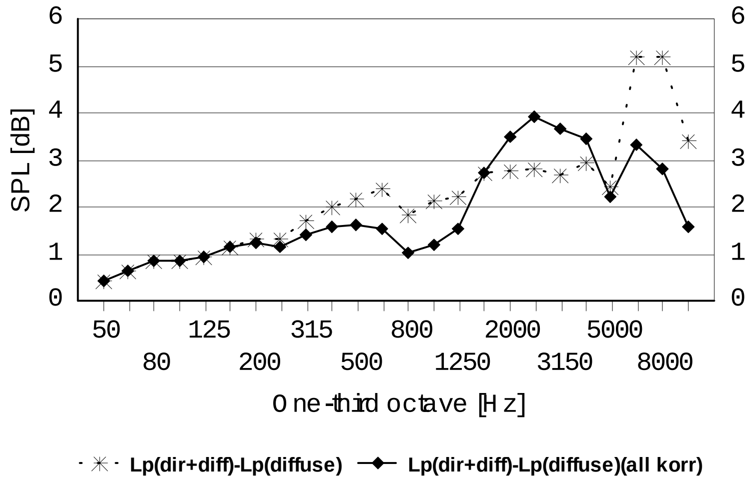

Figure 10 Difference in SPL at the listening position as a function of one-third octaves for the situation where the main lobe of the Quad is facing the listening position (L_p(dir+diff)) and the situation where the Quad is off-set by 90 degrees (L_p(diffuse)). The difference is also shown for a loudness corrected value of the diffuse field (see text).

Signal	Measured SPL difference [dB]	Subjective difference in LDIFF's [dB]
1	3,5	3,4
2	2,2	3,2
3	3,9	3,5
4	3,0	3,7
5	4,6	3,7
6	3,4	3,5
7	3,0	3,2
8	2,3	3,0
9	2,8	3,5

Table 7 Differences in SPL at the listening position for the Quad loudspeaker in the right-front and right-surround channels, respectively together with the corresponding differences in LDIFF values based on the subjective calibrations. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB**.**

Tests of Between-Subjects Effects

Dependent Variable: LDIFF

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	137,010[b]	81	1,691	5,713	,000	462,744	1,000
Intercept	94,263	1	94,263	318,365	,000	318,365	1,000
SIGNAL	12,573	8	1,572	5,308	,000	42,464	,999
CHANNEL	,533	2	,266	,900	,408	1,799	,204
PERSON	73,862	5	14,772	49,893	,000	249,465	1,000
SIGNAL * CHANNE	8,901	16	,556	1,879	,023	30,064	,950
SIGNAL * PERSON	20,660	40	,516	1,744	,006	69,777	,999
CHANNEL * PERSO	20,482	10	2,048	6,918	,000	69,175	1,000
Error	71,652	242	,296				
Total	302,925	324					
Corrected Total	208,662	323					

a. Computed using alpha = ,05

b. R Squared = ,657 (Adjusted R Squared = ,542)

Table 8 ANOVA results for experiment 2. The model only includes significant factors.

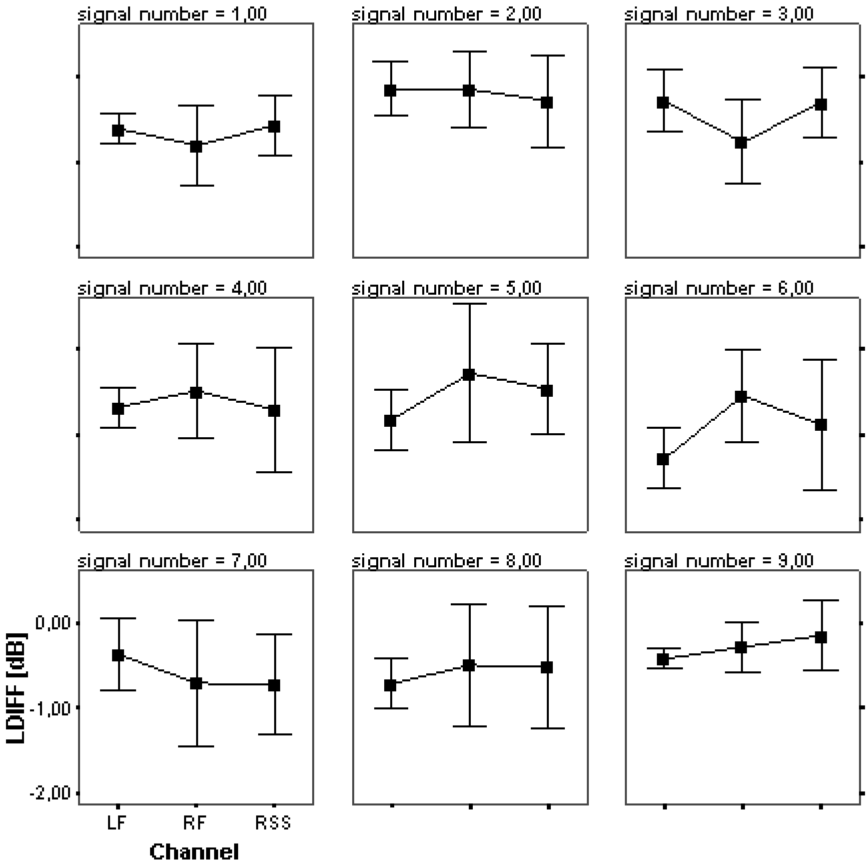

Figure 11 Calibration level for all channels in experiment 2. Shown are mean vaues averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB**.**

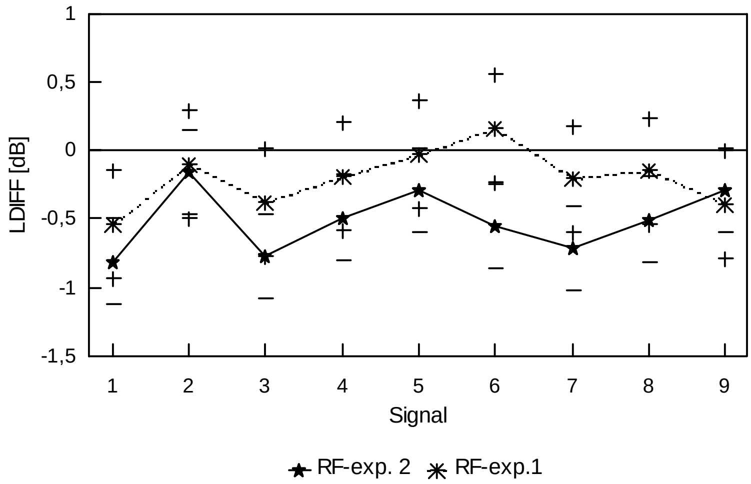

Figure 12 Calibration levels for right-front in experiments 1 and 2, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

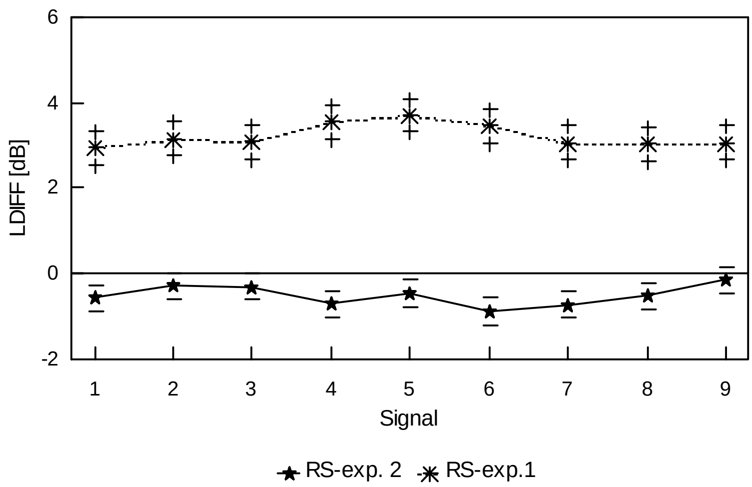

Figure 13 Calibration levels for right-surround in experiments 1 and 2, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

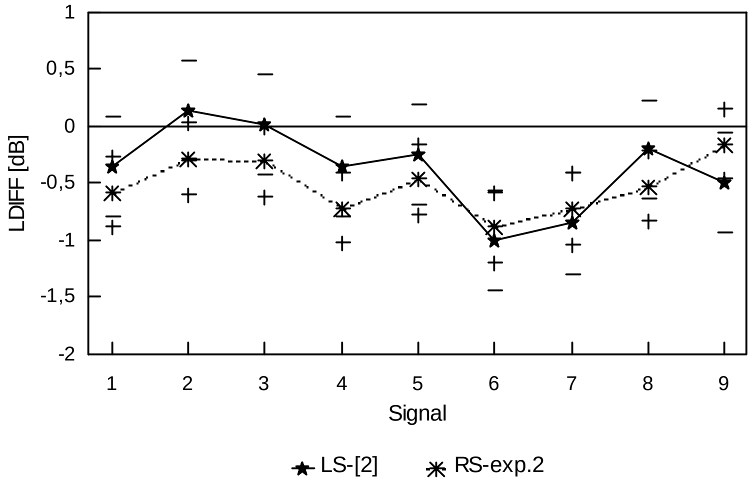

Figure 14 Calibration levels for left-surround in [2] and right-surround in experiment 2, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB**.**

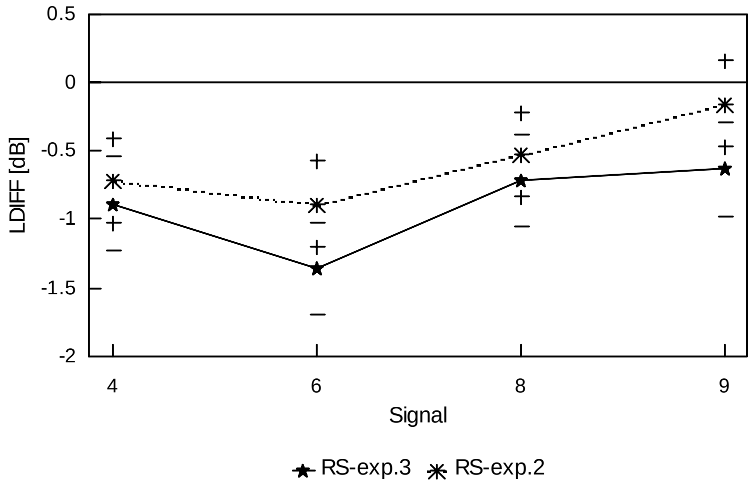

Figure 15 Calibration levels for the right-surround channels in experiments 2 and 3, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

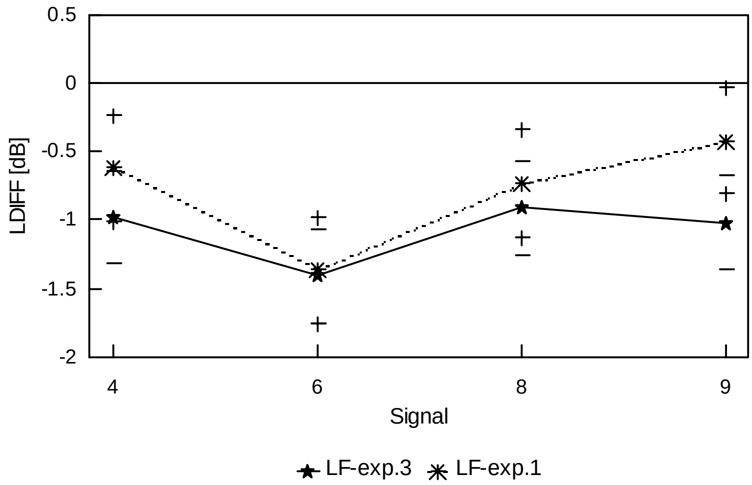

Figure 16 Calibration levels for the left-front channels in experiments 1 and 3, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

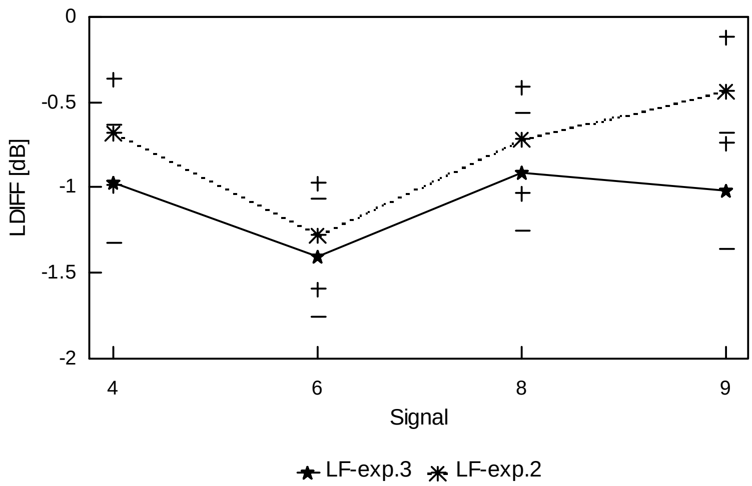

Figure 17 Calibration levels for the left-front channels in experiments 2 and 3, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

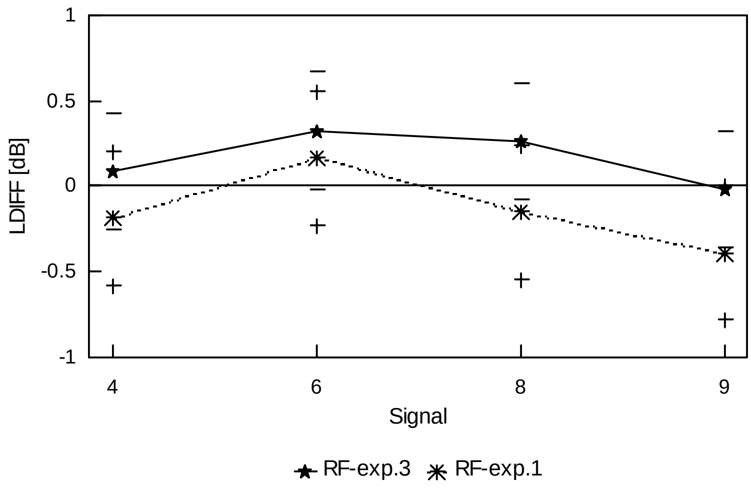

Figure 18 Calibration levels for the right-front channels in experiments 1 and 3, as a function of signal. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

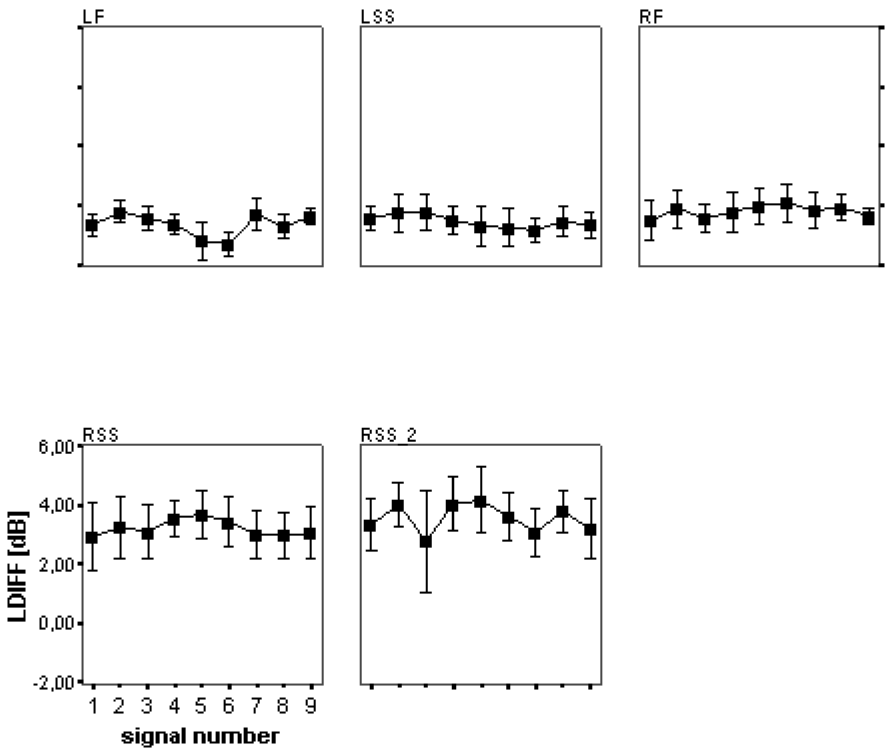

Figure 19 Calibration levels for all channels in experiment 1 shown as a function of signal number. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

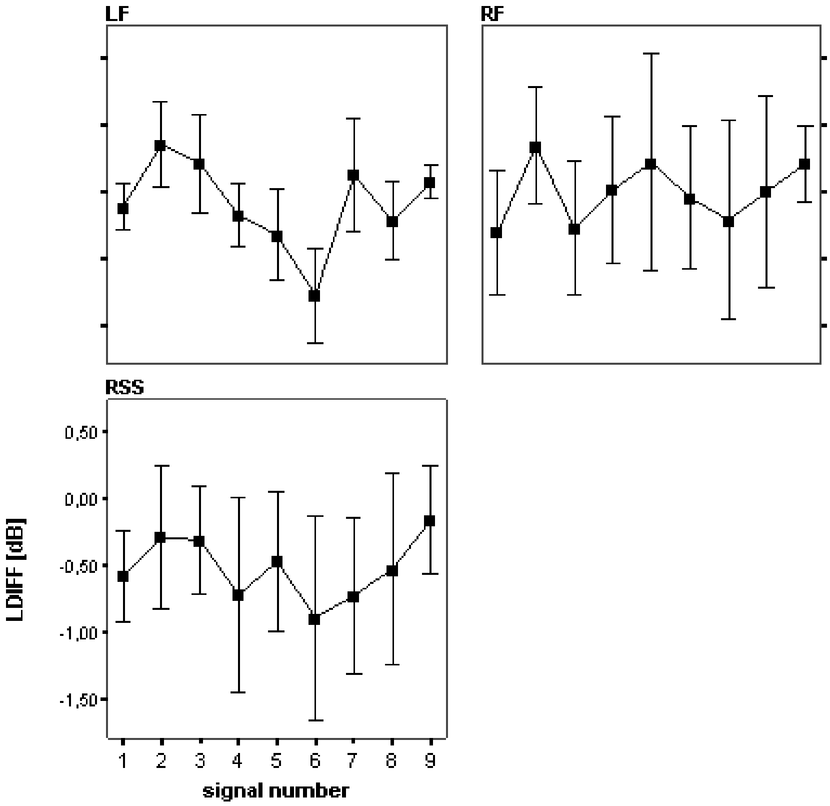

Figure 20 Calibration levels for all channels in experiment 2 shown as a function of signal number. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

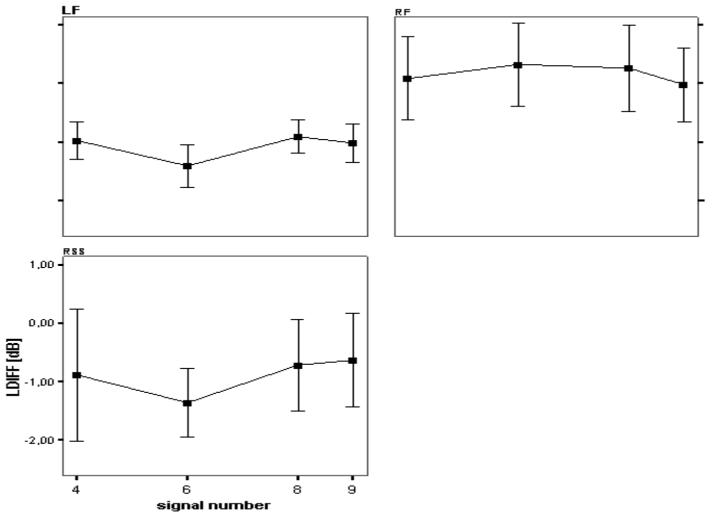

Figure 21 Calibration levels for all channels in experiment 3 shown as a function of signal number. Shown are mean values averaged across subjects and 95% confidence intervals based on the error variance for each signal and channel. LDIFF is the electrical signal level of the adjusted channel relative to the centre channel, and expressed in dB.

Appendix 1. Example of instructions to listeners

Dear listener

You have been asked to take part in a loudness alignment task for multichannel audio systems.

Your task is to adjust the level of the test channel until you are happy that both channels sound equally loud.

For this task you will be asked to compare two sounds. The reference sound is always the centre channel. You can freely switch between the reference and the test channel, which may come from different directions. You will be able to change the level of the test channel using the control system as shown during the instruction phase. When you have done this, the level of the test channel will be noted. You will then be asked to repeat this procedure for several times for different set-up. You have an unlimited time to do this alignment. So just take it easy and only grade the signal when you are happy they are equally loud.

You will be asked to grade XX samples in total, which will be broken into YY session. This should take some 3 or so hours to complete. Remember that there are no correct answers in these tests, so just set the level such that the test and reference signals sound equally loud.

During the course of the session you are asked to remain seated at the centre of the room, and to keep you head facing forward.

You are asked not to discuss the your views on the different noise signals with other listening panel members during the test, though you may do so during the induction and training phase.
Before a test, please ensure that you have not recently been exposed to high sound pressure levels (e.g. a ZZ Top gig). Please inform us if you have had any trouble with you hearing or if you have a cold or the flue. Before a test please ensure that you ears are clean and free from wax, and that you are not hungover.

Good luck and have fun.