Audio Engineering Society

Convention Paper 5368

Presented at the 110th Convention
2001 May 12–15 Amsterdam, The Netherlands

Optimisation and Subjective Assessment of Surround Sound Microphone Arrays

Paul Segar and Francis Rumsey
Institute of Sound Recording, University of Surrey
Guildford, GU2 5XH, UK

ABSTRACT

A number of surround sound arrays have been constructed with closely spaced microphones of cardioid and omni-directional patterns. The spacings and angles between microphones were calculated to test two different psychoacoustic models that aim to provide 360° imaging in a horizontal plane. A series of controlled subjective listening tests have been undertaken, and results are presented comparing image localisation accuracy and localisation confidence between the arrays. Results on the effect of crosstalk between opposite microphones in the arrays are also presented.

0. INTRODUCTION

One of the major advantages of multichannel replay for recordings made in natural acoustics is the ability to convey an enhanced spatial impression of the particular acoustics of the recording venue. Conventional two channel recordings replayed through high-quality loudspeakers at the conventional ±30° positions in front of a listener can portray the texture and timbre of music and localise performers with reasonable accuracy, but are only able to provide a limited spatial representation of the acoustic.

Subjective assessments of the spatial attributes of reproduced sound have been sparsely researched to date [1], possibly because of this difficulty in reproducing a convincing "spatial impression" with two-channel stereo. However, in the field of concert hall acoustics much work has been published over the last 30 years to attempt to establish the key independent, objective parameters which relate to the subjective assessment of an auditorium. Whilst not all of these concert hall parameters are relevant to reproduction in small rooms, they may represent a starting point in seeking to capture an impression of the natural acoustic using multichannel microphone techniques.

It was first recognised by Marshall [2] and Barron [3] that an impression of spaciousness in an auditorium was created by the presence of strong lateral reflections. Spaciousness has since been recognised to be composed of two distinct parts; apparent source width (ASW), and listener envelopment (LEV). Discrete reflections arriving at the listener within a time window from 10ms to about 80ms after the direct sound are not individually distinguishable, but contribute to a broadening of the apparent source width of the direct sound from the performers, tending to draw in the listener. In the time window from 80ms to several hundred milliseconds after the direct sound, the sound energy undergoes an increasing number of reflections becoming essentially reverberant, and contributes to a sense of listener envelopment, particularly if it arrives at the listener from all directions. These effects have been shown to increase with sound intensity, indicating that absolute level of lateral sound energy is important.

In seeking an objective measure of perception of spaciousness, the binaural measurement of inter-aural cross correlation (IACC) was first proposed by Schroeder at al. [4]. Subsequent researchers [5] have considered this measure as separate early and late components, $IACC_E$ corresponding to time windows from 0 to 80ms, and $IACC_L$ from 80ms to beyond 1 second after the direct sound. The early component has shown strong correlation with listeners preferences in concert hall acoustics, indicating that accurate reproduction of early, discrete reflections should be an important consideration in the design of multichannel microphone arrangements.

Results reported in [6] indicated a dependence of the extent of spaciousness on the angle of incidence of the lateral reflections; using pairs of reflections with a time delay of 40ms, a maximum spatial impression was observed at ±90°. Griesinger has also reported the optimum angle for lateral reflections to be ±90° at frequencies below 700Hz, moving progressively towards the median plane at higher frequencies [7]. Ando [8] investigated the preferred direction of arrival of a single reflection with two musical sequences and a time delay of 32ms, comparing subjective preferences with measurements of IACC, and concluded there was a preferred direction for the reflection of approximately 55° ±20°.

In seeking a multichannel microphone arrangement that will provide an optimum spatial impression, there exists a potential trade-off between reproducing the early reflections and the later reverberant sound. To obtain the maximum sense of listener envelopment, it has been proposed that each of the microphone outputs be decorrelated from all others, by spacing them at distances beyond the reverberation radius, and using amplitude panning to provide localisation cues [9]. This should maximise the value of IACC between channels, and hence listener envelopment, but the low correlation between channels does tend to focus the reproduced sound towards the location of the speakers themselves [10]. This approach can however have the benefit of a much less critical listening position, important when multiple listeners are involved.

To enable the discrete, early reflections that produce ASW to be accurately reproduced, a multichannel recording must maintain the time delay and direction of arrival relationships between direct and early reflected sound. This points to an array of microphones that are closely grouped, whose spacings and angles are derived from psychoacoustic models which aim to provide reasonable imaging in a horizontal plane around the listener. The degree of correlation between channels will be higher with this approach, though the typical spacing of microphones still means it will be lower than that obtained by binaural IACC measurement. There will however be high correlation at low frequencies, which should reinforce bass reproduction, a key element in providing good listener envelopment [6]. The listening position with this approach is likely to lead to more of a “sweet spot”, and possibly more critical settings of inter-channel balance.

For commercial recording it is likely that a hybrid of these techniques would be employed, combining a closely spaced array with a number of additional widely spaced microphones.

This paper studies the performance of a number of closely spaced arrays for main microphones, and describes the design and construction of suitable arrays using conventional microphones. It aims to determine from controlled subjective tests whether these arrays can provide stable and accurate localisation of images around the listener. In addition, the effects of crosstalk have been investigated, to determine whether the signals from microphones on the opposite side of the array to the sound source direction played any significant part in assisting or degrading the localisation of the source.

1. ARRAY DESIGN PRINCIPLES

The recording industry is by nature rather conservative, particularly in its use of microphones, and recording engineers have learned from experience which types they prefer in various situations. To be of maximum practical benefit, the aim was to keep the array designs as simple as possible, with the main goals being:

- to make use of conventional microphones
- to feed a single microphone directly to each channel
- to avoid electronic post-processing of the microphone signals in any form (i.e. intensity or delay offsets, or matrixing)

The approach was to investigate whether arrays could be constructed within these constraints, that attempt to provide a seamless 360° horizontal coverage of the sound field. Note however that this does not imply that accurate localisation of images would be possible at all angles around the listener. The array has a definite forward bias, as indeed has the ITU-R BS.775-1 3/2 loudspeaker [11] arrangement intended for reproduction. The large angular spacing of the side and rear loudspeakers of this arrangement may be expected to lead to poor localisation, and producing stable side images is known to be problematic, [12, 13]. Also, there will be geometric distortion between the recording and the reproduction angles of the loudspeakers, to be discussed later.

The types of main microphones commonly used have either cardioid or omni patterns, and both of these were investigated in the multichannel arrays studied.

1.1 Segmentation

To construct the 360° horizontal coverage from an array of five microphones, each pair of microphones was set to cover a designated angular segment, shown in Figure 1. The coverage angle of each segment was designed such that when combined they join one another seamlessly without gaps or overlaps to achieve “critical linking”, following the principles first described in [14,15].

The spacing and angular settings for each microphone pair were based on combinations of intensity and time differences, using well known psychoacoustics for image formation between pairs of loudspeakers, outlined in [16] for two loudspeaker reproduction.

1.2 Coverage angles

It is normal with two-channel cardioid main microphone pairs for the coverage angle to be different to the angular spacing of the microphones themselves; for example an ORTF design has a coverage angle of 95° with microphones angled at 110°. Two-channel pairs also have their microphone angles and their coverage angle symmetrically disposed about the directivity axis .

In the surround array design, each microphone is seen to play a role in two segments. The array is symmetrical about the median plane, meaning that the front centre microphone direction is fixed and in line with the edge of the forward coverage segments, and the rear segment will be symmetrical about its axis of directivity. For the front and lateral segments however, any difference in angle between the microphones and the coverage segment will require the axis of direction of coverage to be different to the axis of microphone direction.

Williams and LeDû [14] have termed this the offset angle, and noted that it can be achieved by either physical position of the microphones to create time or intensity offsets, or by electronic means.

With (ideal) omni microphones, the setting of microphone angle has no influence on the coverage angle of a pair. Thus the coverage angle of omni pairs will always be symmetrically disposed about the directivity axis.

1.3 Microphone positions

The approach used in this paper for setting the positions of the microphones does not rely on the use of time or intensity offsets.

The starting point is first to determine the coverage angle required for each segment, and the microphone angles relative to each edge of the coverage angle, shown if Figure 2. The sum of all coverage segments will be 360°, and the microphone angle relative to the edge of coverage will be used in calculations for two, adjacent segments.

Considering then the situation in the direction of each edge of the coverage segment separately, the intensity differences between the microphones can be calculated. From the psychoacoustics curves of time delay and intensity, the time differences required between the microphones in the direction of each edge of the coverage segment are then found, and hence their separation distances (shown as Δs_1, and Δs_2 in figure 2).

The position of each microphone relative to the other on an (x,y) co-ordinate system is then calculated geometrically from these separation distances and the coverage direction axis.

1.4 Array geometry

In order to achieve critical linkage between the segments, the two constraints of coverage angles and physically realisable geometry have to be met simultaneously. The geometrical constraint requires the microphones to be located at the vertices of a pentangle, determining the relative positions of the microphones.

The use of electronic intensity or time delay differences between microphones was being avoided in these designs, so it was of significance to see whether realisable designs could be achieved by microphone position alone.

1.5 Calculator program

A calculator program was developed following the above design process to generate the microphone positions and angles for arrays of any microphone type. The program allows designs to be generated by an iterative process of varying segment coverage or microphone angles, and has demonstrated that design solutions can indeed be found readily for realistic scenarios.

1.6 Localisation distortion

With two-channel recording, the front coverage angle of a microphone pair is typically 90-100° with cardioid types, and 130-150° with spaced omnis. When replayed through a pair of loudspeakers set at 60° apart, there is clearly considerable angular distortion in the locations of the performers. However, this is acceptable as the microphones are normally placed much closer to the performers than the optimum position of a member of the audience.

In considering a multichannel array, the same situation will apply if the angles of the front segments cover the performers in the same way. However, this will lead to a significant effect on the reproduction of the lateral segment, tending to pull the images at the sides towards the front.

The position of the surround loudspeakers may also exaggerate this effect. The rear coverage angle of these arrays are typically in the range 50-80°, and reproducing this on loudspeakers separated by 140° will also move rear and side images forwards.

This resulting localisation distortion is both an advantage and drawback. The advantage is a tendency to restore the perspective nearer to that of a typical listener in the audience, since reflected sound arriving at the side of the array will be reproduced as coming from a more forward direction, and sound reaching the array from further behind will be reproduced more at the sides. The potential drawback is that little sound will appear to come from behind the listener, and will probably prove difficult to localise.

1.7 Psychoacoustics

Most of the work on the psychoacoustics associated with localisation of images in a reproduction environment has been undertaken with only two loudspeakers, at the normal 60° spacing in front of the listener [17]. The combinations of time delay and intensity to provide localisation entirely in one loudspeaker are well known, represented at the extremes by a time delay difference of about 1.1 ms or intensity difference of approximately 15dB.

It cannot be assumed however that these values of time delay and intensity differences will apply to all loudspeaker pairs in an ITU-R BS775-1 3/2 channel arrangement. A recent study [18] has shown that whilst these values hold true between centre and front loudspeaker pairs, the differences required between the side and rear loudspeaker pairs should be somewhat less, typically 0.9 ms and 12 dB for the lateral segment, and 0.8 ms and 10 dB for the rear segment.

1.8 Crosstalk

In the design of the arrays, the assumption is made that image localisation is created only by one pair of microphones for each segment. Clearly, this is not the case in reality, as sound from any given direction will be picked up by all microphones in the array, and the effect of these “crosstalk” signals must be considered. A degree of intensity reduction will be provided by cardioid microphones pointing away from the sound direction, though omnis will provide very little.

The interactions between these signals is complex and will not be investigated in detail here, though some intuitive observations can be made. In principle, any pair of microphones in the array could attempt to form an image between their respective reproduction loudspeakers. Two situations can be considered; adjacent microphone crosstalk where additional images could be formed by the microphones either side of the pair forming the wanted image, and opposite microphone crosstalk causing delayed sounds to emanate from the opposing loudspeakers.

Adjacent microphone crosstalk is considered in Figure 3, showing the formation of unwanted coverage segments from combinations of one “wanted” microphone and one adjacent “unwanted” microphone. The spacings of these microphone pairs will normally be larger and the microphones more widely angled than those forming the wanted segment, and the resulting coverage segments will therefore be somewhat narrower. The directions of these unwanted segments will also be to either side of the wanted segment, and the degree of overlap will depend on the proximity of the adjacent to the “wanted” microphone. Where overlap does occur, the effect of this crosstalk may be expected to cause a blurring the image localisation to left or right.

Opposite microphone crosstalk may cause microphone pairs on the opposite side of the array to attempt to form images within their back coverage angles. In general though, the time/intensity difference between these signals and those from the “wanted” microphones facing the sound source will be greater than the maximum psychoacoustic values, and the precedence effect should then work to localise the sound source in its true direction.

2. ARRAY DESIGN PROCEDURE

The design procedure for the arrays is fundamentally a two-step process, consisting of determining the front triplet positions based on front coverage requirements, and then deciding the optimum location of the rear microphones for suitable side and rear coverage.

2.1 Front triplet considerations

In order for the musical performers to appear between the front loudspeakers in similar locations to two-channel stereo, the total coverage angle of the front triplet should be very similar to conventional two-channel. Thus each front segment would typically cover 45-55° (90-110° total) with cardioid microphones, or 60-75° (120 - 150° total) with omni microphones. The coverage angle should ideally be determined in practice by the distance from, and size of, the performing group.

The approach to frontal coverage is also important for two-channel compatibility, as it is likely that multi-channel recordings will be mixed down for release on CD and other two-channel formats for the foreseeable future.

Setting the front coverage angle primarily determines the physical width of the array. Selecting a narrow front coverage angle will cause the left and right front microphones to be widely spaced, and vice versa. This also influences the rear coverage angle, as widely spaced front microphones lead to widely spaced rear microphones in these designs, resulting in a narrow rear segment. As noted above, this can be expected to increase localisation distortion.

2.2 Rear microphone considerations

Having set the front microphone positions, it is generally straightforward to determine a solution for the angles and positions of the rear microphones. The calculator program allows the lateral coverage segment and microphone angles to be iteratively increased or decreased until a solution is found.

Typical values of lateral coverage angles are in the range 70-110°, and rear coverage angles in the range 25 - 80°, dependent on front angle settings.

2.3 Different microphones types

All the arrays evaluated here use identical polar patterns for all five microphones, i.e. all cardioid, or all omni. It is however possible to mix types within an array, for example cardioid front triplet with omni rear microphones, and the calculator program can create designs of this type, though the performance of these has not been evaluated at this time.

3. EXPERIMENTAL EVALUATION

A major aim was to investigate whether the arrays would be capable of providing sufficiently accurate localisation, such that the direction and time delay of early reflections contributing to ASW might be preserved in multichannel recording and reproduction.

A total of eight array designs were constructed and evaluated, such that the following variables could be investigated:

- *Cardioid and omni microphones*
  Most reported designs of closely spaced arrays have used cardioid microphones, which use combined intensity/time differences for localisation cues. Whilst this approach is generally considered to provide superior localisation to time difference alone, spaced omnis are often preferred by engineers for their superior low frequency response and sound quality.

- *Front coverage angle*
  As noted above, the choice of front coverage angle affects the side and rear coverage angles of the array significantly, and hence the anticipated localisation distortion.

  Arrays were therefore designed with a choice of narrow and wide front coverage angles to investigate whether the localisation distortion was significant in practice. The cardioid arrays had angles of ±50° and ±65°, and the omni arrays ±55° and ±75°.

- *Psychoacoustic models*
  Both the constant and varying psychoacoustic models of intensity/time delay differences between pairs of microphones were to be investigated. The constant model used the "Williams Curves" values [16] to determine the spacings between all of the microphone pairs, whilst the varying model scaled these curves for the lower values of the side and rear segments.

A total of eight array designs were thus developed, the characteristics and dimensions of which are summarised in Tables 1 to 8.

The predicted angular localisation distortion of reproduction angles versus recording angles for each of these arrays are shown in Table 9.

4. RECORDING PROCESS

The performance of the arrays was to be assessed by a series of controlled subjective listening tests, for which a number of sound sequences were recorded.

In order to minimise the number of variables in the listening tests and to prevent listeners from identifying particular sound sequences, it was decided to replay the sound excerpts over loudspeaker rather than use live sounds. This was necessarily a compromise as the artificiality of this "virtual source" may have been detectable by the listeners and attributed to the arrays themselves. The loudspeaker selected, a Rogers LS3/5A, was a high-quality monitoring unit, with good dispersion characteristics from small diameter bass and treble driver units.

4.1 Programme material

The programme material for the sound sequences was to include examples of musical transients, sounds with slow decay tails, instruments with predominantly high frequency, low frequency, and wide frequency ranges and the human voice. The sources chosen were timpani, piano (high note), piano (low note), castanets and male voice.

Ideally the recording of these sound excerpts would have been made under anechoic conditions, but these facilities were unavailable. The excerpts were therefore recorded close-miked, and (except for piano) in a small booth formed of acoustic screens in Studio 2 at the University of Surrey.

The recordings were recorded in mono with a Neumann KM84 cardioid microphone connected via an SPL Gold pre-amp and phantom supply to a Tascam DA-30 DAT recorder at 16 bit/48 kHz resolution.

4.2 Studio arrangement

The sound excerpts were then replayed through the loudspeaker and recorded using the arrays in Studio 2 at the University of Surrey.

Studio 2 is a medium size, short reverberation studio, intended mostly for pop or jazz music recording, with the following characteristics:

Dimensions:	6.5m (w) x 9.0m (l) x 3.8m (h)
RT60:	c. 0.25 sec
Critical distance:	~2.0m
Description:	Hard walls and ceiling, mostly covered with sound absorbing units, carpeted floor

The microphone arrays were placed approximately centrally in the studio as shown in Figure 4. The height of the array was set at 2.0m.

The Rogers LS3/5A loudspeaker was positioned at a distance of 2.5m from the centre of the array, with the treble unit at a height of 1.6m. Whilst this distance is typical of many recording situations, it was recognised that because of the width of the arrays, this would introduce intensity differences between microphones due to path length of up to 3-5dB at lateral recording angles.

Recordings with the sound excerpts were made with the loudspeaker positioned every 22.5° around each array, a total of 16 positions in total.

4.3 Microphones

- *Cardioid array*
  The cardioid arrays used a set of five Neumann KM84 microphones. To minimise differences in characteristics between the microphones, four of these were from the same batch with consecutive serial numbers. The remaining microphone was used for the centre.

- *Omni array*
  The omni arrays were constructed using a set of five Schoeps MK2S microphones. These were all from the same batch with consecutive serial numbers.

The microphone outputs were fed through SPL Gold pre-amps and recorded as five discrete tracks on a Tascam DA-78 recorder at 16 bit/48 kHz.

5. SUBJECTIVE LISTENING TESTS

5.1 Test design & goals

The subjective listening tests were arranged as three separate tests, such that comparisons could be made between the following different variables involved:

- localisation performance with different microphone types,
- different psychoacoustic models
- different front coverage angles

The first two tests were primarily to assess the localisation performance of each of the arrays, with sound stimuli at all of the 16 positions. The two tests were identical, with Test 1 examining the narrow front coverage segment arrays, and Test 2 the wide front coverage segment arrays.

Four arrays were evaluated in each test, giving a total of 64 excerpts, as follows:

- Cardioid with constant psychoacoustic model
- Cardioid with varying psychoacoustic model
- Omni with constant psychoacoustic model
- Omni with varying psychoacoustic model

Test 3 assessed the effects of opposite microphone crosstalk, for recording angles between 45° and 135° (i.e. 10 locations per array). The technique employed was to run each sound sample twice, once with all five loudspeakers operating, and once with only three loudspeakers active. The three loudspeakers used were the centre, plus the front and rear on the side of the sound source direction. The four array types with constant psychoacoustics were assessed, giving a total of 80 stimuli.

The large number of stimuli involved in each of these tests meant that the subjective assessment had to be limited to the use of a single excerpt. The excerpt selected was a repeated piano note (a C two octaves above middle C). This had a very sharp attack transient and reasonably long decay tail.

The levels of the excerpts from each of the arrays were aligned to within 0.5dB at the 0° recording position, with all channels summed together in a Sonic Solutions digital editor. Whilst this was expected to lead to loudness differences at other recording angles, particularly as the overall intensity of the cardioid array would vary with recording angle, this was considered part of the performance of the arrays and a factor to be investigated.

5.2 Physical set-up

A recently developed software test program, known as ALEX, was used to run the listening test, which took place in the ITU-R BS.1116 [19] listening room at the University of Surrey. The software allowed listeners to control the sound sequences and directly enter their ratings using on-screen sliders, and ran on a Silicon Graphics 02 workstation which also contained all the sound excerpts on hard disc. The multichannel digital audio output from the workstation was via an ADAT interface to a Yamaha 02R mixer, which provided level adjustment and D/A conversion.

Five Genelec 1032A loudspeakers were set up in the standard ITU configuration, at a distance of 2.3m from the listening position. The loudspeakers were level aligned to within 0.1 dBA using a pink noise generator and Brüel and Kjaer 2123 real-time analyser. The absolute intensity of the replay was set at a comfortable level, and was maintained constant for all subjects.

A floor plan of the listening room showing listener and loudspeaker positions is shown in Figure 5.

5.3 Comparison technique

The tests were designed such that the listeners were presented with four stimuli to rate on screen at a time. A blind ABCD paradigm was used, with the four excerpts having being recorded at the same angle from the different arrays, though listeners were not aware of this. The listeners then stepped through either 20 or 24 test screens for all the recording angles. The listener was asked to rate each stimulus individually, and not to compare them in terms of preference.

The array types assigned to A, B, C and D were randomised for each test screen, and the order in which the recording angles were presented was also randomised.

The test sequence was started at one of four different points for different listeners in an attempt to remove any bias due to listeners improving their accuracy through learning at the start of the test, or becoming more random in their scoring through fatigue toward the end of the test.

In order to give the listeners an opportunity to become accustomed to the programme material and the scoring methods, the first four screens of tests were repeated at the end of the test. Listeners were not made aware of this. In the subsequent analysis, the first four sets of test results were then ignored.

5.4 Scales

The listeners were asked to scale four attributes for each of the sound stimuli presented. The attributes were defined as:

- *Image location*
  Listeners were asked to judge the location of the image for each sample, and indicate this on a scale from -180° to +180°. Negative angles were to the left of front centre, and positive to the right.

- *Localisation confidence*
  The confidence in the image position was rated on a scale from 0 to 10. Listeners were asked to rate this based on their confidence in an actual location for the image and its stability (for example, with small head movements), rather than the perceived size of the image.

- *Image shift*
  The amount by which the image shifted between the attack transient and the sound decay tail was rated on a scale of -5 to +5, negative values representing a shift in a counter clockwise direction, and vice versa. This test was aimed at establishing whether the localisation of the transient would appear at a different position to that of the continuous sound of the decay tail.

- *Image distance*
  The apparent distance to the image was rated on a scale from 0 to 10. The directional gain of cardioid microphones normally leads to their placement further away from performers than omni microphones when used as main pairs for two-channel recording. This test, which maintained a constant distance from loudspeaker to microphone array, was aimed at investigating differences in microphone types in surround arrays.

5.5 Test subjects

A total of fifteen listeners took part, split into three groups of five listeners. All were critical and experienced listeners, drawn from staff and final year students on the University of Surrey's Tonmeister undergraduate course.

5.6 Test briefing

The subjects were not aware of the nature of the tests, or the relationship between each of the sound excerpts presented to them. The written instructions to the listeners are shown in Appendix A.

5.7 Test times

It was recognised that the tests would be rather long and complicated, and fatigue may be a problem, even though each listener was asked to participate in only one of the three tests. It was anticipated that each test would take approximately 45 minutes, though in practice test durations varied considerably from 35 to 70 minutes.

6. ANALYSIS OF THE SUBJECTIVE DATA

6.1 Tests 1 and 2: Localisation tests

6.1.1 ANOVA analysis

A multivariate ANOVA analysis was carried out on the data from the five listeners in each of tests 1 and 2. The results using fixed factors of Subject (SUBJNO), Array type (ARRAY), Psychoacoustics (PSYCHO) and Recording angle (RECANGLE), with the dependent variables of Reproduction angle, Error angle, Image confidence, Image shift and Image distance are shown in Tables 10 and 11 respectively. The Error angle was a calculated value, being the difference between the subjective and predicted Reproduction angles given in Table 9. The number of excerpts graded by each listener in each test was identical, giving a balanced design.

From Table 10 on the arrays with narrow front coverage segments, it is noted that there is little dependence on Array type of any variable, except Image distance, and more significance may have been expected from the comparison between cardioid and omni arrays. The Psychoacoustic models also appear to have little significance also on any of the variables. The Reproduction and Error angles are however seen as strongly dependent on the Recording angle, as indeed may be expected. Image confidence is also reasonably dependent on Recording angle.

From the interactions between factors, there is some significance in Reproduction and Error angles on Array type and Recording angle, but no further useful information is revealed.

From Table 11 for Test 2, for wide front coverage angles, much the same dependencies are observed. The Array type shows some significance with Recording angle, Error angle and Image confidence, and a particularly strong significance with Image distance. The Psychoacoustics again show little significance, except with Image distance. The Recording angle again has strong significance on Reproduction and Error angles. Examination of the interaction between factors again reveals little further information of significance.

6.1.2 Reproduction localisation

The dependencies of Reproduction angle on Recording angle for the eight array types in Tests 1 and 2 are shown in Figures 6 to 9, which plot the values of the medians and 25th to 75th percentile bars for the excerpts recorded at 22.5° intervals around the arrays.

Of immediate note is that the median values for all the arrays seem to "flatten off" and appear not to reproduce localised images beyond about 120°. To an extent, this performance is anticipated by the predicted localisation given in Table 9, though it also probably points to an inability to form convincing images behind the position of the rear loudspeaker positions.

All the arrays appear to have some difficulty in localising images at 180°, which often appeared at or close to 0° instead, showing in the figures as very large error bars. There was a similar though lesser effect also with localisation at 0°, which was occasionally confused for 180°. In both cases, localisation was however always close to the median plane. It is interesting to note that this effect occurred almost equally with cardioid and omni microphones.

From the length of the error bars, there appears also to be more uncertainty in the localisation of images between the side loudspeakers, at reproduction angles between 30° and 110°. This is consistent with the difficulty of forming stable side images reported in [13] and [18]. Interestingly, the errors tend to decrease as the image becomes localised more in the rear loudspeaker, and then increase again for greater angles.

In comparing the two psychoacoustic models in (a) and (b) of each figure, it is noted that there is considerable similarity in their localisation performance. Since the positions of the front triplet are the same for both models, the performance at smaller angles should be similar. The closer spacings of the rear microphones for the varying psychoacoustics may have been expected to give a lower slope to the angular variation across the lateral segment coverage angle, though any effect due to this is not apparent in these results and a closer study of this area is probably necessary to identify any differences. The only array to show a noticeable difference between the two psychoacoustic models is the omni array with narrow front coverage, shown in figure 8, and the differences here are mostly at the smaller angles.

The effect of the front coverage angle on the array localisation performance is seen to be very significant. Comparing firstly the two cardioid designs in Figures 6 and 7, the narrow front coverage (±50°) demonstrates a reasonably linear relationship between recording and reproduction angles. The array with wide coverage angle (±65°) however shows a very different story with a very non-linear relationship between recording and reproduction angles, with reproduction angles being substantially less than recording angles for smaller values, then rising steeply to reach the maximum reproduction angle at 90° recording angle and flattening out at larger angles.

Similar behaviour is seen with the omni arrays in Figures 8 and 9. The omni array with the narrow front coverage angle (±55°) in Figure 8 exhibits somewhat exaggerated reproduction angles at smaller angles, greater even than the corresponding recording angles There also appears to be a greater uncertainty in the frontal image localisation than with the other arrays, judging by the length of the error bars. The reason for this is probably adjacent microphone crosstalk, due to the close proximity of the front and rear microphones with this array design.

By way of contrast, the wide front coverage (±75°) omni array shown in Figure 9 displays very accurate localisation at the smaller angles, comparable to the wide cardioid array. The overall shape of the characteristic is reasonably linear to 112° recording angle , then flattening off at larger angles.

It would appear therefore from these results that the choice of front coverage angle has a significant effect on the localisation performance of the array, and indeed optimum values would appear to exist to give the best angular linearities.

6.1.3 *Error in localisation angles*

As a measure of the localisation performance of each of the arrays against their predicted behaviour, plots of Error angle against Recording angle are shown in Figures 10 to 13. The Error angle is the difference between the measured and predicted Reproduction angle, and a positive error indicates that the reproduced image was localised at a greater angle, i.e. further towards the rear, than its predicted position.

All of the arrays exhibit a similar general characteristic, with a positive angular error of about 40-60° at 80-100° recording angle, then decreasing to zero error around 160° recording angle and going negative beyond this. The wide front coverage arrays appear to exhibit about 20° greater maximum errors at the sides with a more peaked characteristic, the range of angles for positive errors corresponding closely to the lateral segment coverage angles. Whilst the reasons for these errors is unclear, they do therefore appear to be associated with the localisation abilities of the side pairs of microphones.

Three of the four arrays have small errors at the 22.5° and 45° recording angles, indicating they are behaving close to predicted. The omni array with narrow front segment is the exception, clearly showing significant errors which pull the front imaging further to the side, probably due to adjacent microphone crosstalk noted earlier.

Examining the psychoacoustic models, only small differences exist between the two graphs for each array. For the omni array with narrow front coverage, the varying psychoacoustic model shows greater errors, which because of the closer proximity of front and rear microphones indicates the presence of adjacent microphone crosstalk.

6.1.4 *Image confidence, Image shift and Image distance*

The results of image confidence, image shift and image distance for the arrays are shown in Figures 14 and 15.

One of the main points of interest in these experiments was to compare the localisation abilities of cardioid and omni arrays. Omnis, relying on time differences alone are generally considered to produce more diffuse and less stable images than cardioids. The results here shown in figures 14(a) and 15(a) indicate that although the mean values of image confidence are slightly lower for the omni arrays, the differences are small and not statistically significant.

The image shift between the transient and sound decay tail was close to zero for all the arrays, though this may have been difficult for the subjects to judge in combination with the other factors, leading them to indicate little difference. During the tests, some listeners had reported unusual image shift effects with some of the excerpts, though most listeners noted little observable shift.

The perception of image distance is shown in figures 14(c) and 15(c), and indicates that the omni arrays appeared to provide localisation slightly closer than cardioid arrays. This was probably due to loudness differences between omni and cardioid arrays, and possibly also timbral differences between the microphone types.

6.2 Crosstalk : Test 3

6.2.1 *ANOVA analysis*

A multivariate ANOVA analysis was carried out on the data from the five listeners in Test 3. The results using fixed factors of Subject (SUBJNO), Array type (ARRAY), Number of loudspeakers (SPEAKERS) and Recording angle (RECANGLE), with Reproduction angle and Image confidence for the dependent variables are shown in Tables 12. The number of excerpts graded by each listener was identical, giving a balanced design.

The table indicates a strong dependence of reproduction angle on Recording angle, with only small dependence on the number of replay loudspeakers. The Image confidence is seen to show some dependence on the number of loudspeakers.

6.2.2 *Image confidence*

The mean values and associated 95% confidence intervals for image confidence of each of the arrays when reproduced through three or five loudspeakers are shown in Figure 16. Each of the four arrays displays some loss of image confidence when all five loudspeakers are used, though not statistically significant except for the wide front coverage omni array, due probably to the small number of subjects involved.

6.2.3 *Difference angle*

To assess whether the localisation at each position altered between when all five or only three loudspeakers were used, a difference angle was calculated between the resulting subjective localisation values. Figure 17 shows this angular difference for each array type, which indicates very little change in localisation. The omni array with wide front coverage again appears to provide the greatest difference, though not significant statistically with the number of subjects involved, which may be due to the narrower width of this array design.

7. INFORMAL SUBJECTIVE ASSESSMENTS

A number of experimental recordings were made in collaboration with the BBC during the 2000 season of Promenade Concerts in the Royal Albert Hall, London. These used an array of omni microphones, with a front coverage angle set at ±60°, and versions with constant and varying psychoacoustics were tried in various locations in the hall throughout the season.

Informal views expressed by professional balancers that have heard these recordings indicate that a good sense of the original acoustic is conveyed, which seems to extend seamlessly to either side of the front left and right loudspeakers.

For one of the Promenade concerts, there was the opportunity to compare a recording made on the omni array with a BBC multi-microphone recording [20]. The multi-microphone recording used a combination of spaced omni main microphones and spot microphones for the performers, with a total of five omni microphones placed well back in the auditorium to provide decorrelated sound in the surround and front channels, to which additional electronic reverberation was also added.

The two approaches produced rather different results. The multi-microphone recording produced a sense of being present in a large and reverberant acoustic, and the listening position was relatively uncritical to obtaining good frontal localisation. The microphone array gave a much more intimate impression and sense of natural acoustic, considered to be similar to that of a member of the audience at a reasonable distance from the performers. The array was also free from phasing effects observed on applause with the multi-microphone recording when the listener moved their head forward and backward.

Side-to-side listening position was relatively uncritical, though front-to-back position was more sensitive with the array.

8. CONCLUSIONS

The work described in this paper has demonstrated that closely spaced microphone arrays can be constructed for five channel surround recording using conventional microphones. The design approach has segmented the horizontal 360° plane into five coverage segments, which critically link together without gaps or overlap. The coverage of each segment is then assigned to a pair of microphones, whose positions are calculated based on combined intensity/time differences.

A software calculator program has been developed, which has been used to design the eight arrays using cardioid and omni microphones evaluated here. The calculator has shown that solutions for array designs can be found for most combinations of front coverage angles of cardioid or omni microphones, without the need for electronic intensity or time delays between microphones to achieve critical linkage between segments.

The localisation performance of the arrays has been assessed by recording sound extracts at 22.5° intervals around each array in a studio with short reverberation time, and judging the results by controlled subjective listening tests. Because of the large number of excerpts and variables judged during the listening tests, only five subjects took part in each of the three tests, and this has necessarily has reduced the accuracy and increased the statistical confidence intervals for some of the results. The listening tests were also rather long, which may have also introduced a degree of inaccuracy due to listener fatigue.

From the subjective results obtained, both cardioid and omni microphones have demonstrated they can produce localisation of images with similar degrees of confidence. Arrays constructed from both types of microphone have demonstrated their ability to localise images up to positions approximately ±120° either side of the listener, beyond which, because of the localisation distortion introduced by the array design, it was predicted that images would not be formed. For sounds directly behind the listener, and to a lesser extent with sounds at the centre front, there was some front-to-back confusion of image location, though the subjective localisation was always close to the median plane. Whether this is a concern in practice is uncertain, but does point to some criticality in level and/or delay settings between front and rear speakers. Front centre localisation could in practice probably be improved by deliberately compromising the rear performance, by delaying or reducing the level in the rear channels.

There was some evidence that localisation of stable images between the side loudspeakers was less certain, as has been noted previously by other researchers. However, stable images were formed at around 90-112° recording angles, when localisation was in the rear loudspeaker direction. With the two wide front coverage arrays (which also had wide rear coverage segments), the localisation again became more uncertain at greater angles.

The choice of front coverage angle was found to be significant for both cardioid and omni arrays, both types demonstrating an optimum angle to achieve good angular linearity between recording and reproduction angle. For cardioid microphones this appeared to be close to the narrow front coverage array angle of ±50°, and for omnis it appeared somewhere between the two arrays, probably around ±65-70°.

Comparison of the subjective localisation performance of the arrays with predicted values shows that they all consistently produced images farther towards the rear at recording angles between 60° and 160°. The wide front coverage arrays displayed a greater maximum error of about 60°, compared to the maximum 40° of the narrow front coverage arrays. All arrays displayed these localisation errors over the predicted angular range of their lateral segments. The reasons for these errors are unclear, as other researchers [18] evaluating side image performance, albeit with only two loudspeakers, have not reported any evidence of this effect. There was some evidence of adjacent microphone crosstalk effects between front and rear microphones when these became too close, as with the narrow front coverage omni array, causing frontal image localisation to be shifted rearwards

The two psychoacoustic models did not show significant differences in localisation performances, indicating that the positioning of the rear microphones was less critical. The closer spacing between front and rear microphones with the varying psychoacoustic designs did however appear to exaggerate probable adjacent microphone crosstalk with the narrow front coverage omni array design.

In terms of image confidence there was little difference between cardioid and omni arrays, with omnis showing only a very small and statistically insignificant lower confidence level. There was also no evidence of image shifting between the attack transient of the excerpt and the decaying sound tail. Perceived image distance was closer for the omni arrays, probably due to loudness or timbral differences.

The effect of crosstalk between microphones on opposite sides of the arrays did not appear in the subjective tests to have a significant effect in the localisation performance, even with omni microphones. There did though appear to be a consistent, though small, loss of image confidence when using five rather than just three loudspeakers.

The results of the subjective tests reported here have necessarily been limited to recordings made in a single non-reverberant acoustic, with a sound source at a fixed distance and a single sound extract. Whilst this arrangement was selected to minimise the influence of room reflections, and hence provide a good environment for assessing localisation performance, it does not represent the type of acoustic we are ultimately attempting to portray using these arrays for music recordings. Further work is therefore needed to assess the array performances in typical recording acoustics with sound sources at different distances, where in addition to localisation performance, aspects such as ASW, listener envelopment and naturalness could be judged.

In summary therefore the use of closely-spaced microphone arrays as main microphones, correctly optimised for angular linearity through selection of front coverage angle, do appear capable of producing convincing localisation over an arc of ±120° in front of the listener when using a five-channel ITU loudspeaker arrangement. Such arrays should therefore prove capable of conveying the angular and time delay relationships of reflected sounds from the recording acoustic to the listening environment, and practical recordings made to date have shown promising results in capturing the essential characteristics of the original acoustic.

9. ACKNOWLEDGEMENTS

The authors would like to thank the staff and students at University of Surrey who participated in the subjective listening tests, Russell Mason for his advice during the preparation and analysis of the listening test material, and Andrew Thompson for technical support throughout the project. The authors are very grateful also to Protape Ltd for the loan of the Schoeps microphones, and to the BBC for the opportunity to make recordings during the Proms 2000. Particular thanks go to the BBC studio managers and outside broadcast staff for their help and support with technical facilities during the season.

10. REFERENCES

1. Rumsey, F., 1998; “Subjective Assessment of the Spatial Attributes of Reproduced Sound”, Proc. 15th International AES Conference, Copenhagen Denmark, pp122-135

2. Marshall A.H., 1968; “Acoustical Determinants for the Architectural Design of Concert Halls”, Arch. Sci. Rev., Australia 11, 81-87
3. Barron M., 1971; “The Subjective Effects of First Reflections in Concert Halls - The need for Lateral Reflections”, J. Sound Vib, 15, 475-494
4. Schroeder M.R., Gottlob D. and Siebrasse K.F., 1974; “Comparative Study of European Concert Halls; Correlation of Subjective Preferences with Geometric and Acoustic Parameters”, Journal of the Acoustical Society of America, Vol 56, no. 4 (October), pp. 1195 1201.
5. Beranek L.L ,1996; “Concert and Opera Halls How They Sound” , Acoustical Society of America, ISBN 1-56396-530-5
6. Barron M. and Marshall A,H., 1981; “Spatial Impression due to Early Lateral Reflections in Concert Halls”, J. Sound Vib, 77, 211-232
7. Griesinger D., 1998; “General Overview of Spatial Impression, Envelopment, Localisation and Externalisation”, Proc 15th Int Conf AES on Small Room Acoustics, pp. 136-149, Denmark Oct/Nov 1998
8. Ando Y., 1985; “Concert Hall Acoustics”, Springer Verlag, Berlin
9. Griesinger D., 2000; “The Theory and Practice of Perceptual Modelling - How to use Electronic Reverberation to ad Depth and Envelopment without Reducing Clarity”, Tonmeister Conference, Hannover, November 2000
10. Theile G., 2000; “Multichannel Natural Recording based on Psychoacoustic Principles”; 108th AES Convention, Paris
11. ITU-R BS 775-1, 1992-4; “Multichannel Stereophonic Sound System With and Without Accompanying Picture”; Geneva
12. Ratcliffe P.A., 1974; “Properties of Hearing related to Quadraphonic Reproduction”, BBC Report RD 38.
13. Thiele G., Plenge G., 1976; “Localisation of Phantom Sources”, J. Audio Eng Soc. (25), pp. 196-200
14. Williams M., Le Du G,1999; “Microphone Array Analysis for Multichannel Sound Recording”; 107th AES Convention, New York, Sept. 1999
15. Williams M., Le Du G,2000; “Multichannel Microphone Array Design”, 108th AES Convention, Paris, February 2000
16. Williams M., 1987; “Unified Theory of Microphone Systems for Stereophonic Sound Recording”; 82nd AES Convention, Preprint 4997
17. Blauert J., 1996, “Spatial Hearing - The Pychophysics of Human Sound Localisation”, MIT Press, Cambridge, Mass., ISBN 0-262 02413-6
18. Martin,W. Woszczyk W., Corey J., Quesnel R., 1999; “Sound Source Localisation in a Five-Channel Surround Sound Reproduction System”; 107th AES Convention, New York
19. ITU-R BS 1116, 1994; “Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems”, Geneva
20. Hancock S., 2000; “The Proms in Surround”, LineUp, Journal of the Institute of Broadcast Sound, No. 76, pp. 16-20, Dec 2000/ Jan 2001.

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	50	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	50	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**70.0**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	114	**Psychoacoustics**	
Lateral mic angle *(deg)*	114	Max intensity diff (dB)	14.89
Lateral mic spacing *(cm)*	**9.4**	Max time diff. (ms)	1.13
Rear coverage segment			
Rear coverage angle *(deg)*	32	**Psychoacoustics**	
Rear mic angles *(deg)*	32	Max intensity diff (dB)	14.89
Rear mic spacing *(cm)*	124.8	Max time diff. (ms)	1.13
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-63.4	29.6	-50
Right front	63.4	29.6	50
Left surround	-60.7	38.6	-164
Right surround	60.7	38.6	164

Table 1. Cardioid array design with narrow front coverage angle and constant psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	50	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	50	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**70.0**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	118	**Psychoacoustics**	
Lateral mic angle *(deg)*	118	Max intensity diff (dB)	12.00
Lateral mic spacing *(cm)*	**1.5**	Max time diff. (ms)	0.90
Rear coverage segment			
Rear coverage angle *(deg)*	24	**Psychoacoustics**	
Rear mic angles *(deg)*	24	Max intensity diff (dB)	10.00
Rear mic spacing *(cm)*	121.6	Max time diff. (ms)	0.80
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-63.4	29.6	-50
Right front	63.4	29.6	50
Left surround	-63.0	31.0	-168
Right surround	63.0	31.0	168

Table 2. Cardioid array design with narrow front coverage angle and varying psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	65	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	64	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**47.2**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	88	**Psychoacoustics**	
Lateral mic angle *(deg)*	88	Max intensity diff (dB)	14.89
Lateral mic spacing *(cm)*	**26.5**	Max time diff. (ms)	1.13
Rear coverage segment			
Rear coverage angle *(deg)*	54	**Psychoacoustics**	
Rear mic angles *(deg)*	56	Max intensity diff (dB)	14.89
Rear mic spacing *(cm)*	62.3	Max time diff. (ms)	1.13
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-39.8	25.3	-64
Right front	39.8	25.3	64
Left surround	-30.7	50.3	-152
Right surround	30.7	50.3	152

Table 3. Cardioid array design with wide front coverage angle and constant psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	65	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	65	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**47.2**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	96	**Psychoacoustics**	
Lateral mic angle *(deg)*	96	Max intensity diff (dB)	12.00
Lateral mic spacing *(cm)*	**14.1**	Max time diff. (ms)	0.90
Rear coverage segment			
Rear coverage angle *(deg)*	38	**Psychoacoustics**	
Rear mic angles *(deg)*	38	Max intensity diff (dB)	10.00
Rear mic spacing *(cm)*	69.4	Max time diff. (ms)	0.80
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-39.8	25.3	-65
Right front	39.8	25.3	65
Left surround	-34.3	38.3	-161
Right surround	34.3	38.3	161

Table 4. Cardioid array design with wide front coverage angle and varying psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	55	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	60	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**84.2**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	106	**Psychoacoustics**	
Lateral mic angle *(deg)*	75	Max intensity diff (dB)	14.89
Lateral mic spacing *(cm)*	**48.7**	Max time diff. (ms)	1.13
Rear coverage segment			
Rear coverage angle *(deg)*	38	**Psychoacoustics**	
Rear mic angles *(deg)*	90	Max intensity diff (dB)	14.89
Rear mic spacing *(cm)*	119.4	Max time diff. (ms)	1.13
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-74.6	38.9	-55
Right front	74.6	38.9	55
Left surround	-59.6	85.1	-161
Right surround	59.6	85.1	161

Table 5. Omni array design with narrow front coverage angle and constant psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	55	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	60	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**84.2**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	112	**Psychoacoustics**	
Lateral mic angle *(deg)*	75	Max intensity diff (dB)	12.00
Lateral mic spacing *(cm)*	**37.3**	Max time diff. (ms)	0.90
Rear coverage segment			
Rear coverage angle *(deg)*	26	**Psychoacoustics**	
Rear mic angles *(deg)*	90	Max intensity diff (dB)	10.00
Rear mic spacing *(cm)*	122.3	Max time diff. (ms)	0.80
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-74.6	38.9	-55
Right front	74.6	38.9	55
Left surround	-61.3	73.7	-167
Right surround	61.3	73.7	167

Table 6. Omni array design with narrow front coverage angle and varying psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	75	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	60	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**63.8**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	64	**Psychoacoustics**	
Lateral mic angle *(deg)*	75	Max intensity diff (dB)	14.89
Lateral mic spacing *(cm)*	**73.3**	Max time diff. (ms)	1.13
Rear coverage segment			
Rear coverage angle *(deg)*	82	**Psychoacoustics**	
Rear mic angles *(deg)*	90	Max intensity diff (dB)	14.89
Rear mic spacing *(cm)*	59.2	Max time diff. (ms)	1.13
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-50.6	38.9	-75
Right front	50.6	38.9	75
Left surround	-29.2	109.0	-139
Right surround	29.2	109.0	139

Table 7. Omni array design with wide front coverage angle and constant psychoacoustics

Front coverage segment			
Front centre to L or R coverage angle *(deg)*	75	**Psychoacoustics**	
Front centre to L or R pair angle (deg)	60	Max intensity diff (dB)	14.89
Front centre to L or R spacing *(cm)*	**63.8**	Max time diff. (ms)	1.13
Lateral coverage segment			
Lateral coverage angle *(deg)*	78	**Psychoacoustics**	
Lateral mic angle *(deg)*	75	Max intensity diff (dB)	12.00
Lateral mic spacing *(cm)*	**49.2**	Max time diff. (ms)	0.90
Rear coverage segment			
Rear coverage angle *(deg)*	54	**Psychoacoustics**	
Rear mic angles *(deg)*	90	Max intensity diff (dB)	10.00
Rear mic spacing *(cm)*	60.6	Max time diff. (ms)	0.80
Coordinates			
(NB All reference front centre)	**X *(cm)***	**Y *(cm)***	**Angle *(deg)***
Centre Front	0.0	0.0	0
Left front	-50.6	38.9	-45
Right front	50.6	38.9	45
Left surround	-30.6	83.8	-180
Right surround	30.6	83.8	180

Table 8. Omni array design with wide front coverage angle and varying psychoacoustics

Recording angle	**0°**	**22.5°**	**45°**	**67.5°**	**90°**	**112.5°**	**135°**	**157.5°**	**180°**
Cardioid arrays									
Narrow front coverage with constant psychoacoustics	0°	13.5°	27.0°	42.3°	58.1°	73.9°	89.6°	105.4°	180°
Narrow front coverage with varying psychoacoustics	0°	13.5°	27.0°	41.9°	57.1°	72.4°	87.6°	102.9°	180°
Wide front coverage with constant psychoacoustics	0°	10.4°	20.8°	32.3°	52.7°	73.2°	93.6°	121.7°	180°
Wide front coverage with varying psychoacoustics	0°	10.4°	20.8°	32.1°	50.8°	69.6°	88.3°	107.1°	180°
Omni arrays									
Narrow front coverage with constant psychoacoustics	0°	12.3°	24.5°	39.4°	56.4°	73.4°	90.4°	107.4°	180°
Narrow front coverage with varying psychoacoustics	0°	12.3°	24.5°	38.9°	55.0°	71.1°	87.1°	103.2°	180°
Wide front coverage with constant psychoacoustics	0°	9.0°	18.0°	27.0°	48.8°	76.9°	105.0°	141.6°	180°
Wide front coverage with varying psychoacoustics	0°	9.0°	18.0°	27.0°	45.4°	68.5°	91.5°	121.7°	180°

Table 9. Predicted reproduction angles for sources recorded at 22.5° intervals with the eight array types

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
SUBJNO	Reproduction angle	148722.531	4	37180.633	49.527	.000
	Error angle	149081.760	4	37270.440	46.225	.000
	Image confidence	385.263	4	96.316	54.814	.000
	Image shift	11.511	4	2.878	3.458	.010
	Image distance	128.021	4	32.005	16.595	.000
ARRAY	Reproduction angle	346.391	1	346.391	.461	.498
	Error angle	104.423	1	104.423	.130	.719
	Image confidence	7.124	1	7.124	4.054	.046
	Image shift	1.611	1	1.611	1.936	.166
	Image distance	127.992	1	127.992	66.366	.000
PSYCHO	Reproduction angle	656.335	1	656.335	.874	.351
	Error angle	1157.708	1	1157.708	1.436	.233
	Image confidence	.303	1	.303	.172	.679
	Image shift	.315	1	.315	.379	.539
	Image distance	2.623	1	2.623	1.360	.246
RECANGLE	Reproduction angle	188959.430	8	23619.929	31.463	.000
	Error angle	204681.411	8	25585.176	31.732	.000
	Image confidence	173.075	8	21.634	12.312	.000
	Image shift	10.620	8	1.328	1.595	.131
	Image distance	20.494	8	2.562	1.328	.234
SUBJNO * ARRAY	Reproduction angle	6159.925	4	1539.981	2.051	.090
	Error angle	6238.538	4	1559.635	1.934	.108
	Image confidence	21.680	4	5.420	3.085	.018
	Image shift	1.312	4	.328	.394	.813
	Image distance	2.405	4	.601	.312	.870
SUBJNO * PSYCHO	Reproduction angle	778.375	4	194.594	.259	.904
	Error angle	815.618	4	203.905	.253	.907
	Image confidence	4.562	4	1.140	.649	.628
	Image shift	1.345	4	.336	.404	.805
	Image distance	8.212	4	2.053	1.064	.377
ARRAY * PSYCHO	Reproduction angle	591.968	1	591.968	.789	.376
	Error angle	731.934	1	731.934	.908	.342
	Image confidence	1.130	1	1.130	.643	.424
	Image shift	.163	1	.163	.196	.659
	Image distance	4.089	1	4.089	2.120	.148
SUBJNO * RECANGLE	Reproduction angle	149409.321	32	4669.041	6.219	.000
	Error angle	150188.724	32	4693.398	5.821	.000
	Image confidence	240.393	32	7.512	4.275	.000
	Image shift	38.508	32	1.203	1.446	.076
	Image distance	64.348	32	2.011	1.043	.417
ARRAY * RECANGLE	Reproduction angle	54406.679	8	6800.835	9.059	.000
	Error angle	57415.735	8	7176.967	8.901	.000
	Image confidence	33.764	8	4.221	2.402	.019
	Image shift	10.181	8	1.273	1.529	.152
	Image distance	44.932	8	5.616	2.912	.005
PSYCHO * RECANGLE	Reproduction angle	6578.023	8	822.253	1.095	.370
	Error angle	6489.476	8	811.184	1.006	.434
	Image confidence	19.202	8	2.400	1.366	.217
	Image shift	5.803	8	.725	.872	.542
	Image distance	1.988	8	.248	.129	.998

Table 10. Test 1: Multivariate ANOVA results table for all listeners, with subject, array, psychoacoustics and recording angle as fixed factors, and reproduction angle, error angle image confidence, Image shift and image distance as dependent variables

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
SUBJNO	Reproduction angle	127667.841	4	31916.960	43.948	.000
	Error angle	79142.841	4	19785.710	27.244	.000
	Image confidence	210.932	4	52.733	20.089	.000
	Image shift	13.991	4	3.498	1.047	.385
	Image distance	372.636	4	93.159	62.703	.000
ARRAY	Reproduction angle	4978.182	1	4978.182	6.855	.010
	Error angle	4203.636	1	4203.636	5.788	.017
	Image confidence	13.475	1	13.475	5.133	.025
	Image shift	.511	1	.511	.153	.696
	Image distance	268.945	1	268.945	181.021	.000
PSYCHO	Reproduction angle	712.727	1	712.727	.981	.324
	Error angle	32.727	1	32.727	.045	.832
	Image confidence	3.111	1	3.111	1.185	.278
	Image shift	2.184	1	2.184	.654	.420
	Image distance	23.645	1	23.645	15.915	.000
RECANGLE	Reproduction angle	365513.047	8	45689.131	62.911	.000
	Error angle	203856.172	8	25482.021	35.087	.000
	Image confidence	81.172	8	10.146	3.865	.000
	Image shift	7.497	8	.937	.281	.971
	Image distance	11.772	8	1.471	.990	.446
SUBJNO * ARRAY	Reproduction angle	10079.205	4	2519.801	3.470	.010
	Error angle	20305.114	4	5076.278	6.990	.000
	Image confidence	22.423	4	5.606	2.135	.080
	Image shift	1.045	4	.261	.078	.989
	Image distance	100.418	4	25.105	16.897	.000
SUBJNO * PSYCHO	Reproduction angle	2351.477	4	587.869	.809	.521
	Error angle	10794.659	4	2698.665	3.716	.007
	Image confidence	14.059	4	3.515	1.339	.259
	Image shift	8.736	4	2.184	.654	.625
	Image distance	15.855	4	3.964	2.668	.035
ARRAY * PSYCHO	Reproduction angle	460.227	1	460.227	.634	.427
	Error angle	311.136	1	311.136	.428	.514
	Image confidence	2.184	1	2.184	.832	.363
	Image shift	5.457	1	5.457	1.634	.203
	Image distance	18.409	1	18.409	12.391	.001
SUBJNO * RECANGLE	Reproduction angle	167877.969	32	5246.187	7.224	.000
	Error angle	128352.969	32	4011.030	5.523	.000
	Image confidence	154.313	32	4.822	1.837	.009
	Image shift	52.175	32	1.630	.488	.990
	Image distance	85.869	32	2.683	1.806	.010
ARRAY * RECANGLE	Reproduction angle	16864.922	8	2108.115	2.903	.005
	Error angle	27694.297	8	3461.787	4.767	.000
	Image confidence	41.022	8	5.128	1.953	.057
	Image shift	10.622	8	1.328	.398	.920
	Image distance	15.722	8	1.965	1.323	.237
PSYCHO * RECANGLE	Reproduction angle	2343.047	8	292.881	.403	.917
	Error angle	28524.922	8	3565.615	4.910	.000
	Image confidence	22.122	8	2.765	1.053	.399
	Image shift	20.547	8	2.568	.769	.630
	Image distance	7.297	8	.912	.614	.765

Table 11. Test 2: Multivariate ANOVA results table for all listeners, with subject, array, psychoacoustics and recording angle as fixed factors, and reproduction angle, error angle image confidence, Image shift and image distance as dependent variables

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
SUBJNO	Reproduction angle	100290.250	4	25072.563	47.820	.000
	Image confidence	53.665	4	13.416	3.861	.005
ARRAY	Reproduction angle	6698.187	3	2232.729	4.258	.006
	Image confidence	36.890	3	12.297	3.539	.016
RECANGLE	Reproduction angle	195644.625	4	48911.156	93.286	.000
	Image confidence	31.815	4	7.954	2.289	.061
SPEAKERS	Reproduction angle	915.063	1	915.063	1.745	.188
	Image confidence	94.090	1	94.090	27.076	.000
SUBJNO * ARRAY	Reproduction angle	14190.250	12	1182.521	2.255	.011
	Image confidence	31.135	12	2.595	.747	.704
SUBJNO * RECANGLE	Reproduction angle	76352.875	16	4772.055	9.102	.000
	Image confidence	140.085	16	8.755	2.520	.002
ARRAY * RECANGLE	Reproduction angle	10603.375	12	883.615	1.685	.072
	Image confidence	100.085	12	8.340	2.400	.006
SUBJNO * SPEAKERS	Reproduction angle	6532.750	4	1633.187	3.115	.016
	Image confidence	2.885	4	.721	.208	.934
ARRAY * SPEAKERS	Reproduction angle	944.187	3	314.729	.600	.616
	Image confidence	19.490	3	6.497	1.870	.136
RECANGLE * SPEAKERS	Reproduction angle	850.875	4	212.719	.406	.804
	Image confidence	21.935	4	5.484	1.578	.182

Table 12. Crosstalk Test 3: Multivariate ANOVA results table for all listeners, with subject, array, recording angle and number of replay loudspeakers as fixed factors, and reproduction angle and image confidence as dependent variables

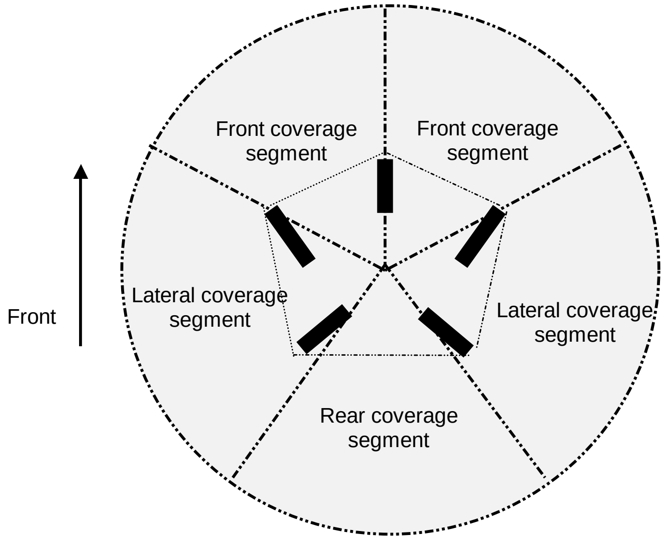


Figure 1. Plan view of array, showing general arrangement of microphones and coverage segments

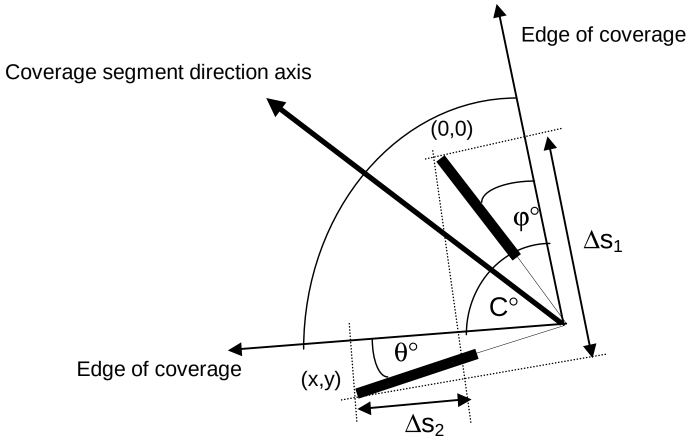


Figure 2. Relationship of microphone positions to coverage segment size and direction axis

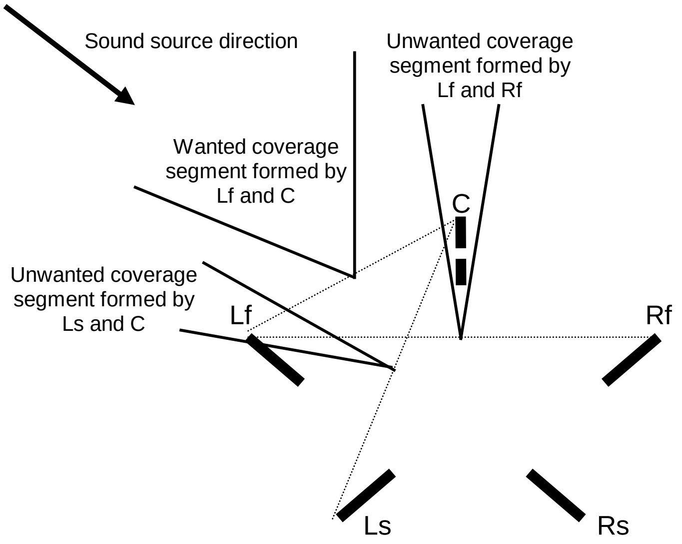


Figure 3. Effect of adjacent microphone crosstalk in the formation of unwanted images

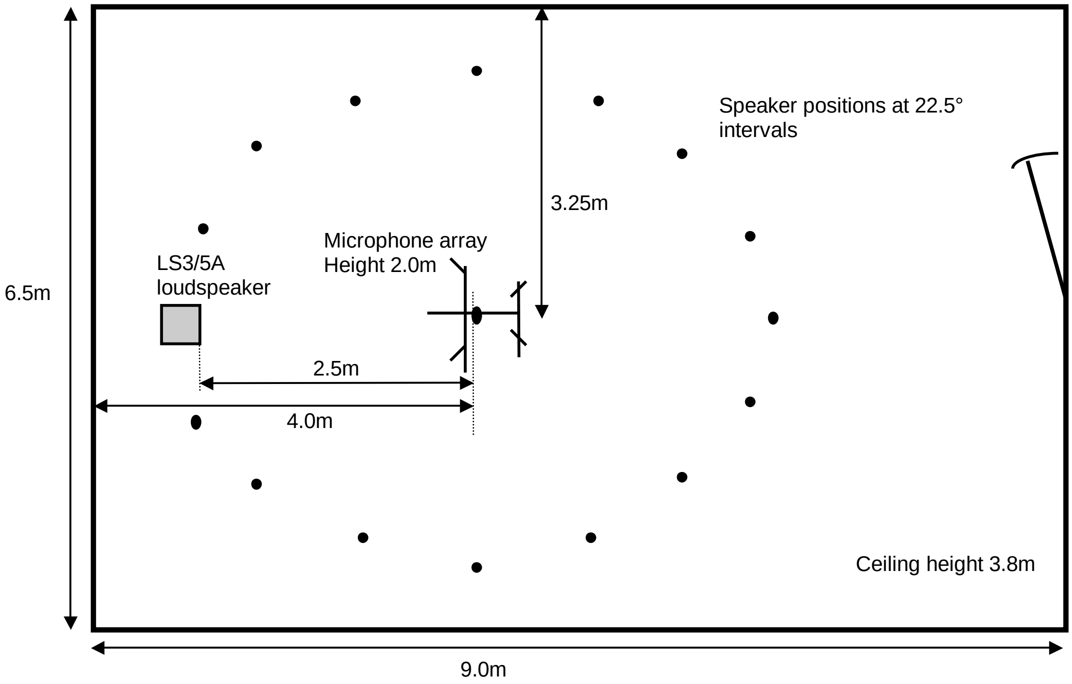


Figure 4. Plan view of Studio 2 recording arrangement

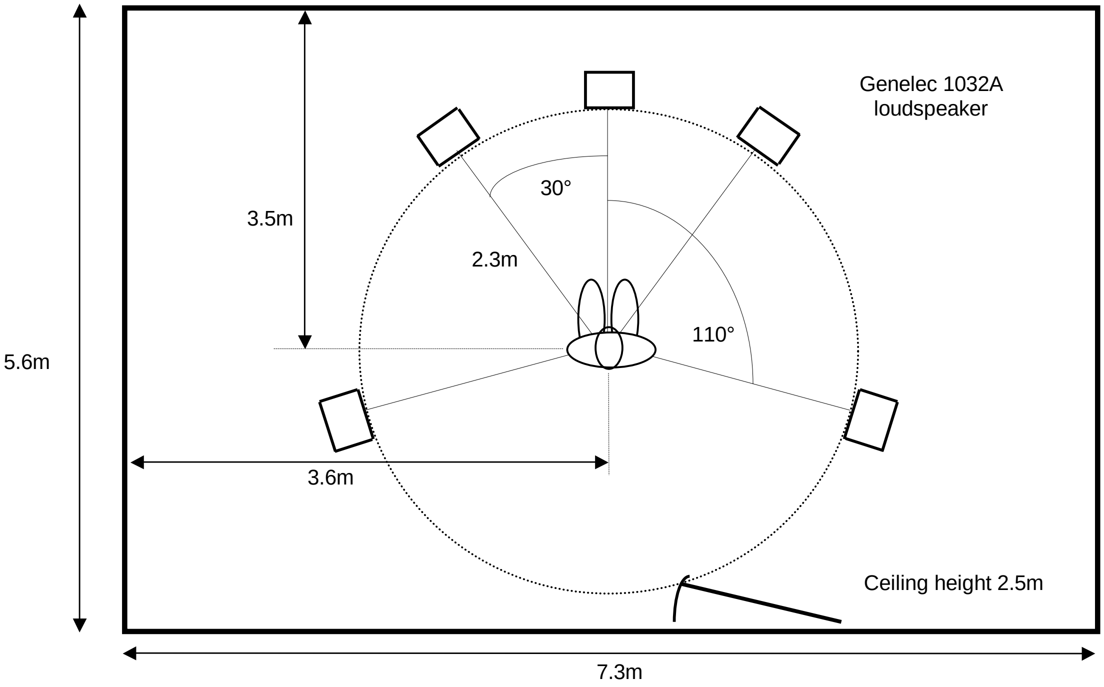
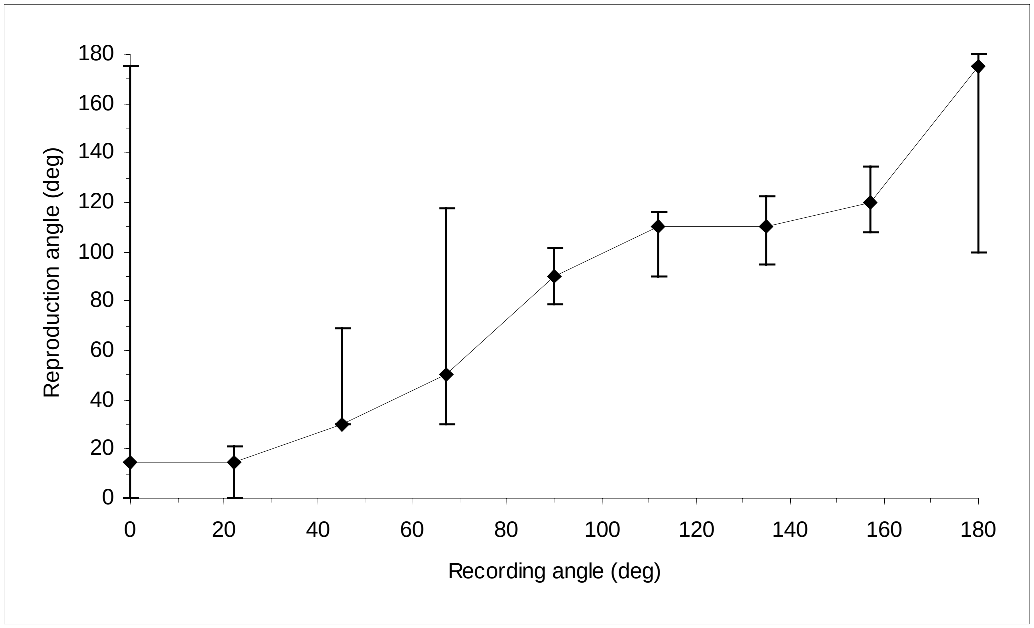
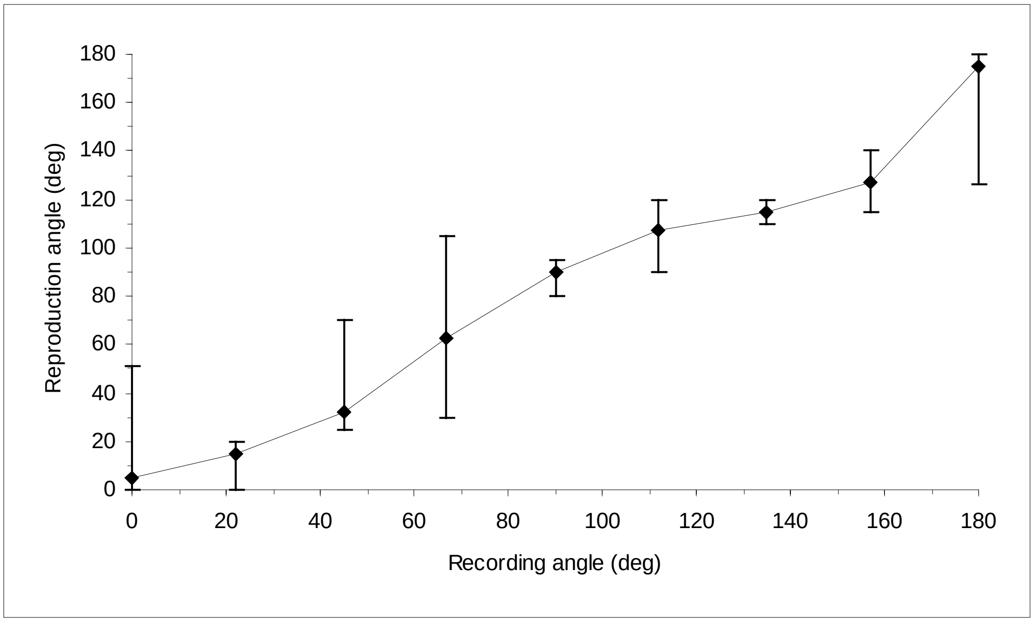


Figure 5. Plan view of listening room floor plan

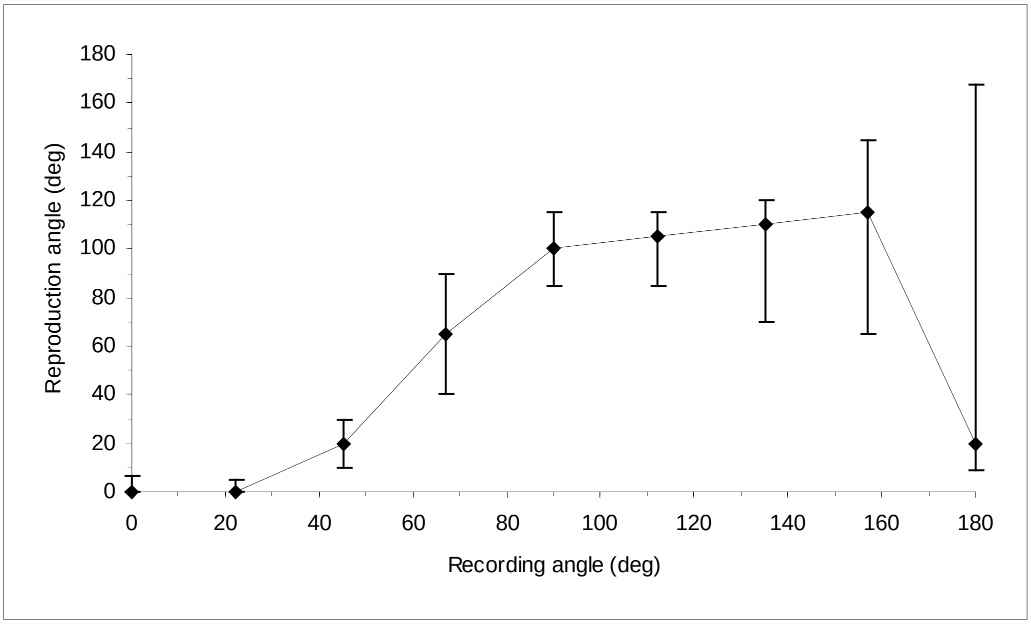


(a) Array with constant psychoacoustics

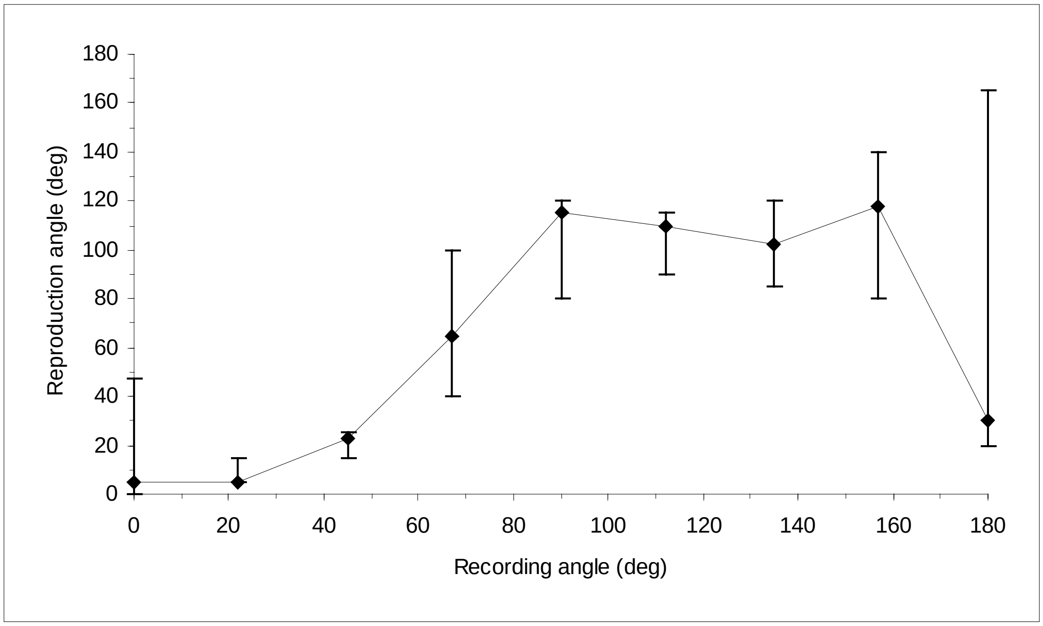


(b) Array with varying psychoacoustics

Figure 6: Test 1: Cardioid arrays with narrow front coverage angle: Median values and associated 25th to 75th percentile of reproduction angle vs. recording angle

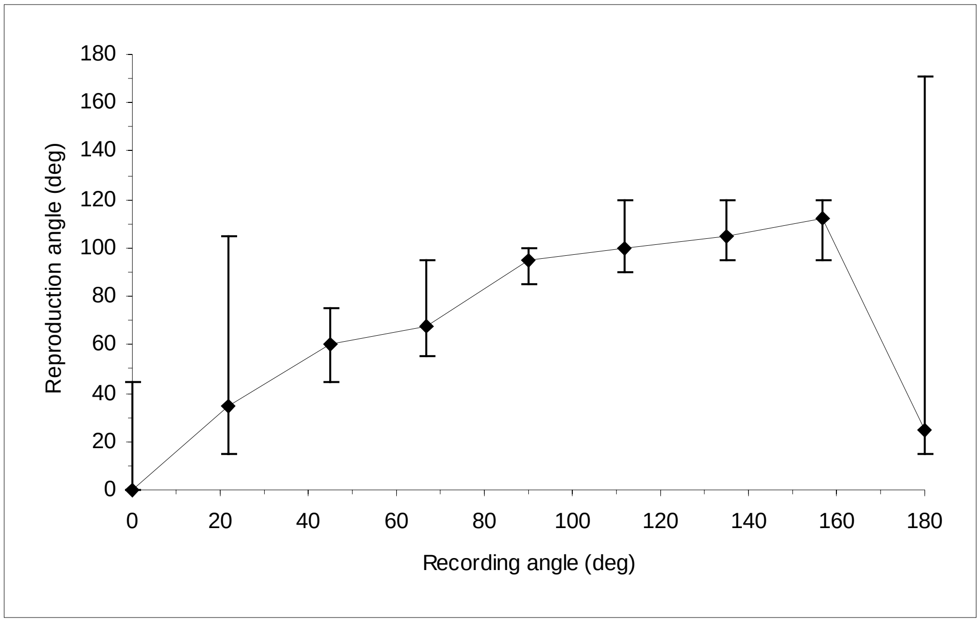


(a) Array with constant psychoacoustics

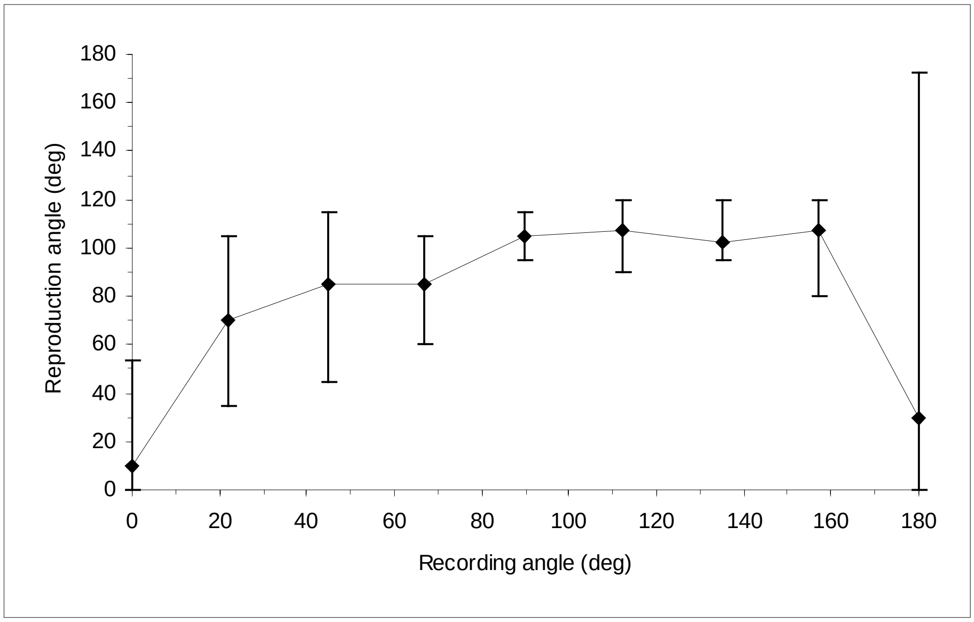


(b) Array with varying psychoacoustics

Figure 7: Test 2: Cardioid arrays with wide front coverage angle: Median values and associated 25th to 75th percentile of reproduction angle vs. recording angle

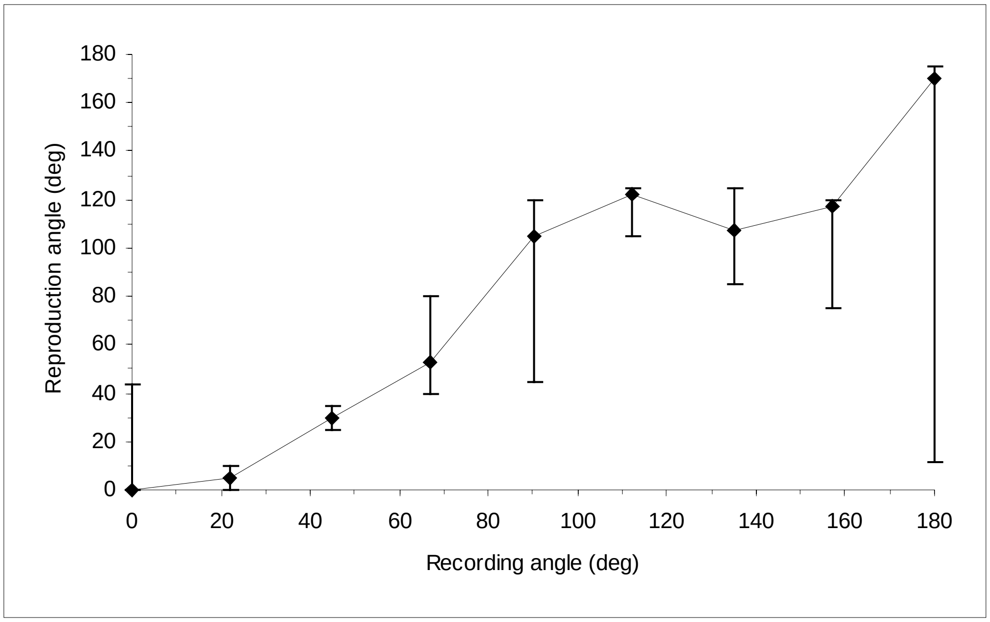


(a) Array with constant psychoacoustics

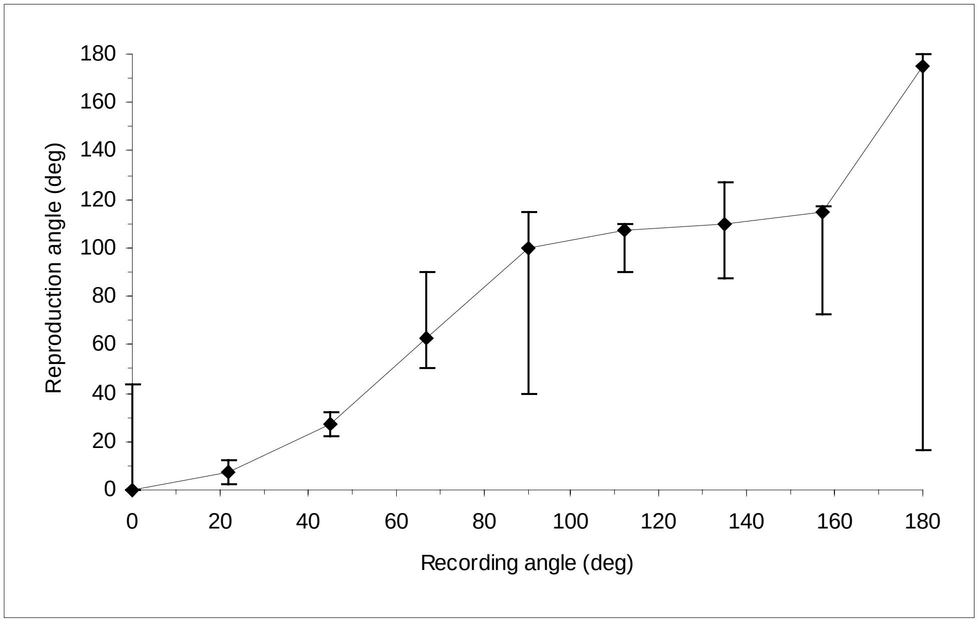


(b) Array with varying psychoacoustics

Figure 8: Test 1: Omni arrays with narrow front coverage angle: Median values and associated 25th to 75th percentile of reproduction angle vs. recording angle

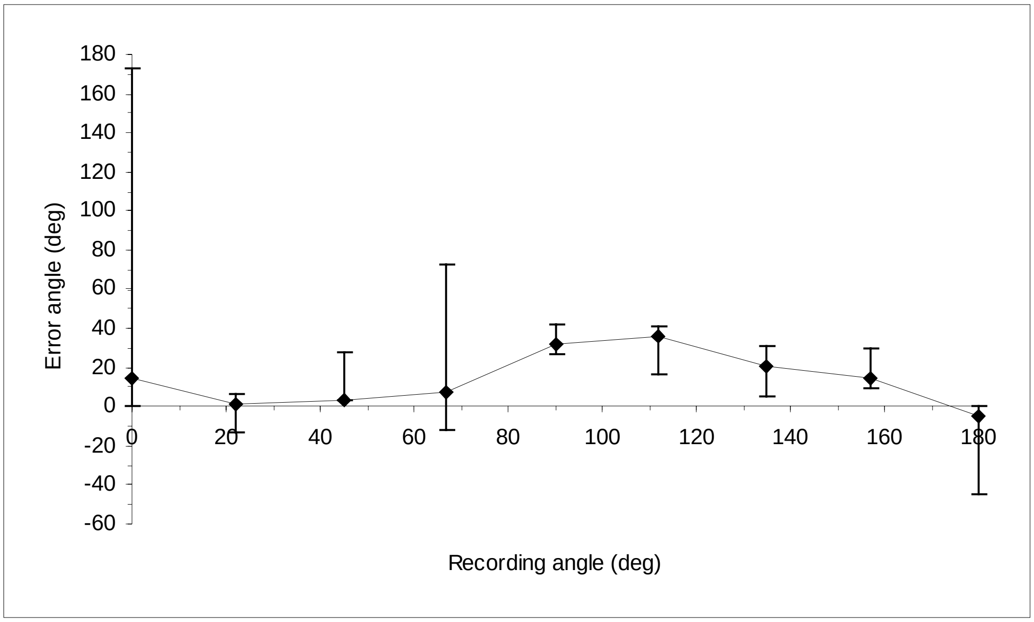


(a) Array with constant psychoacoustics

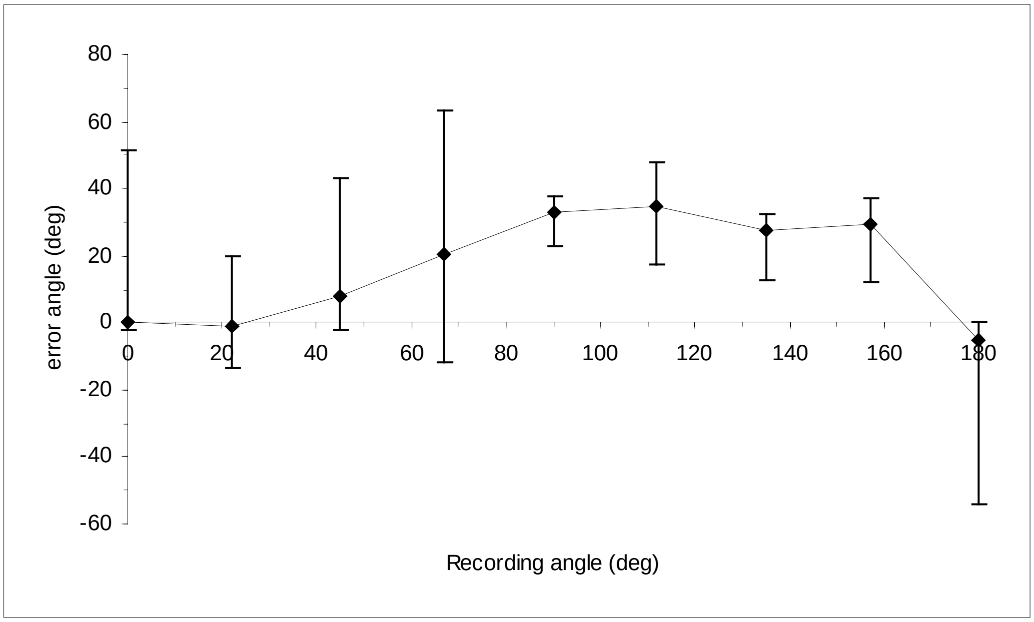


(b) Array with varying psychoacoustics

Figure 9: Test 2: Omni arrays with wide front coverage angle: Median values and associated 25th to 75th percentile of reproduction angle vs. recording angle

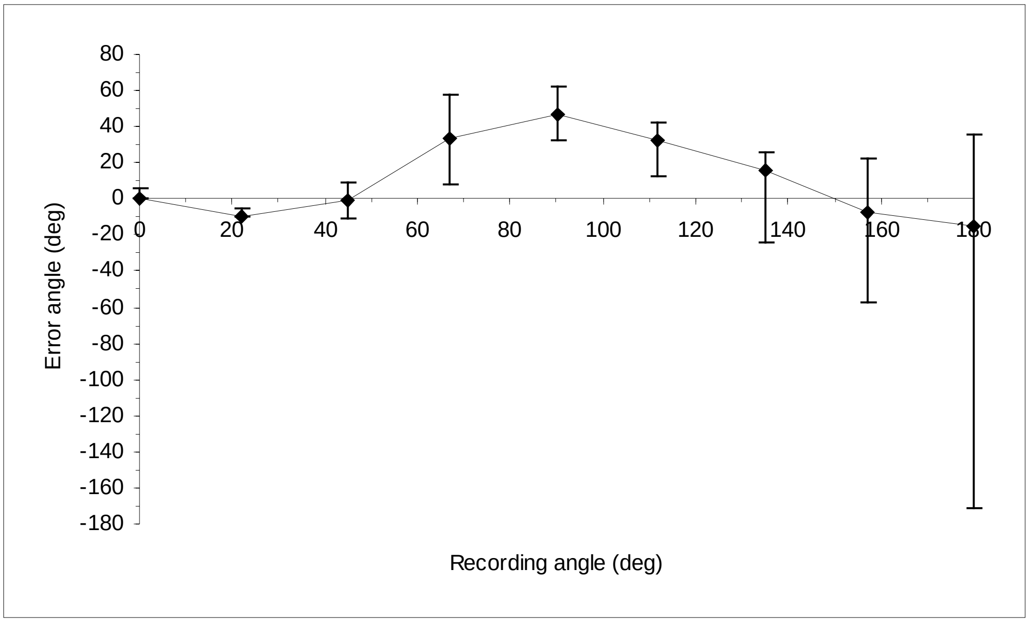


(a) Array with constant psychoacoustics

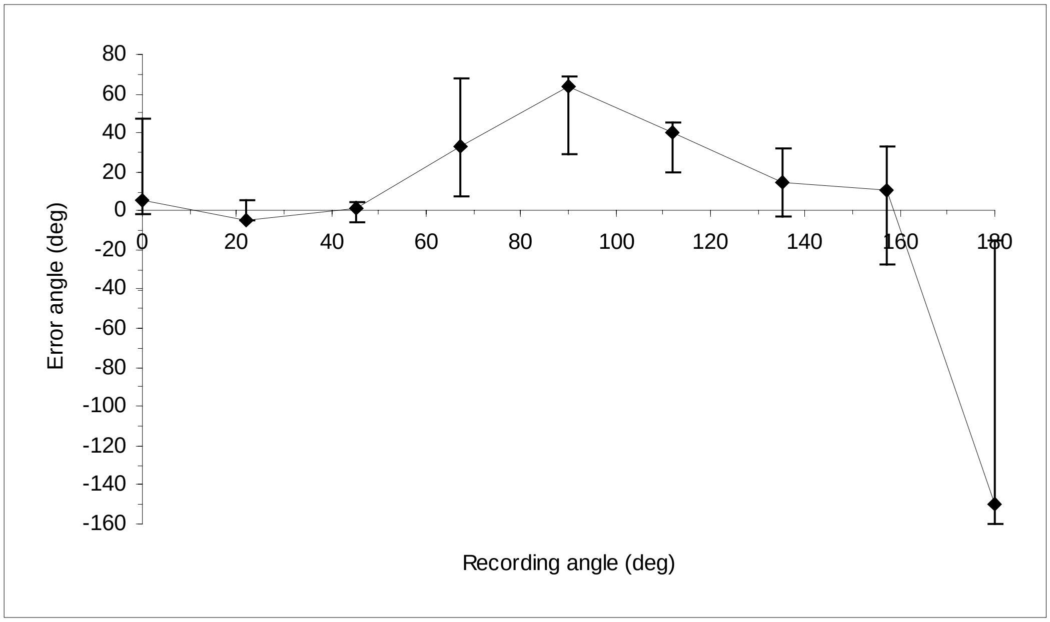


(b) Array with varying psychoacoustics

Figure 10: Test 1: Cardioid arrays with narrow front coverage angle: Median values and associated 25th to 75th percentile of error angle vs. recording angle

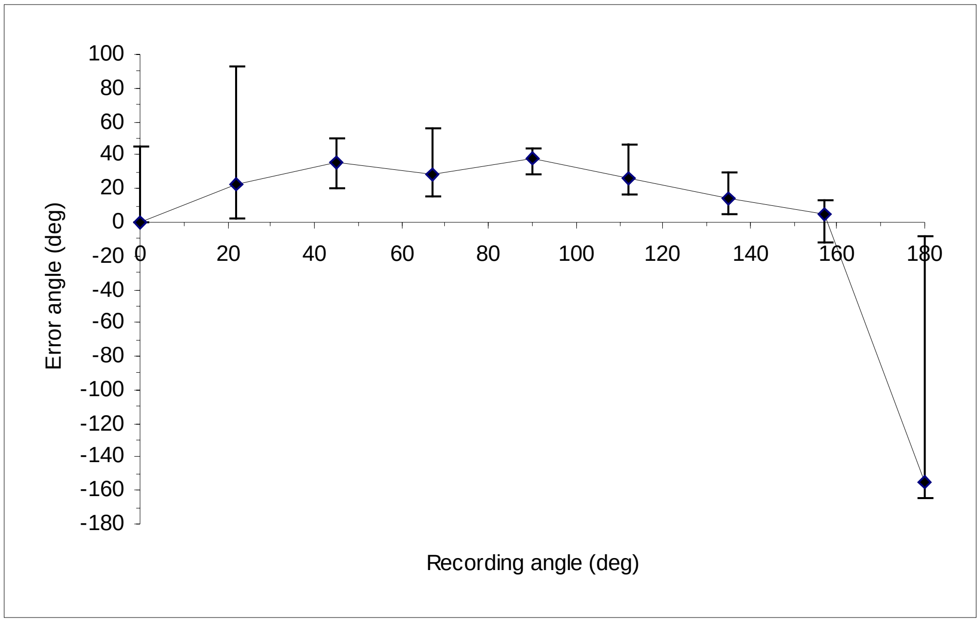


(a) Array with constant psychoacoustics

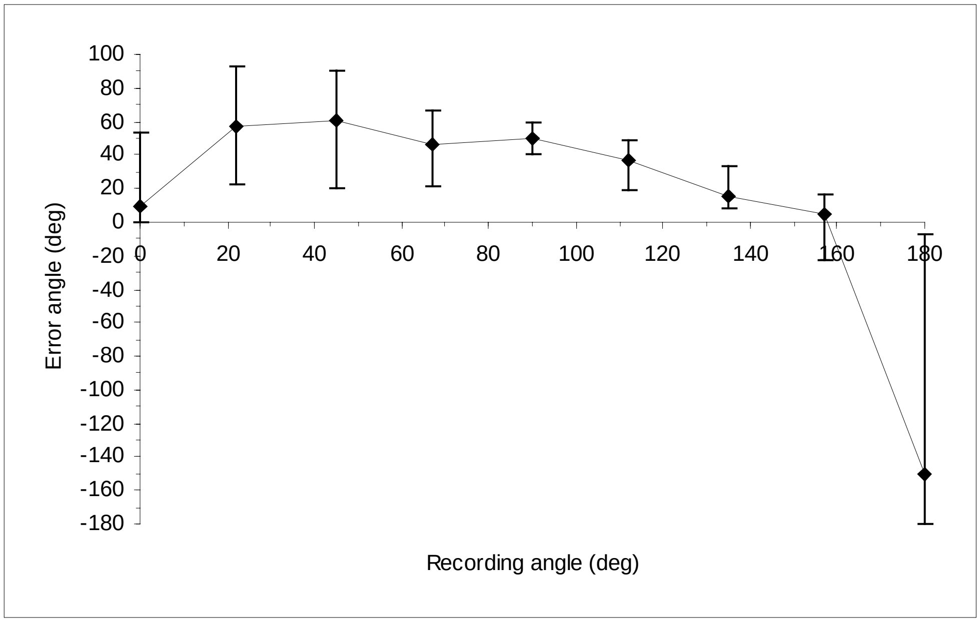


(b) Array with varying psychoacoustics

Figure 11: Test 2: Cardioid arrays with wide front coverage angle: Median values and associated 25th to 75th percentile of error angle vs. recording angle

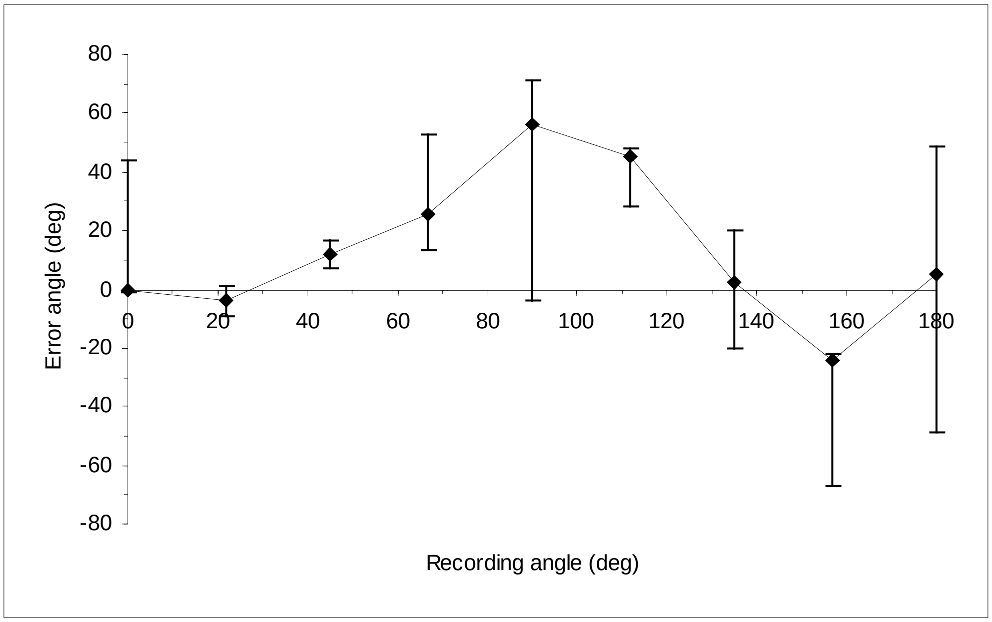


(a) Array with constant psychoacoustics

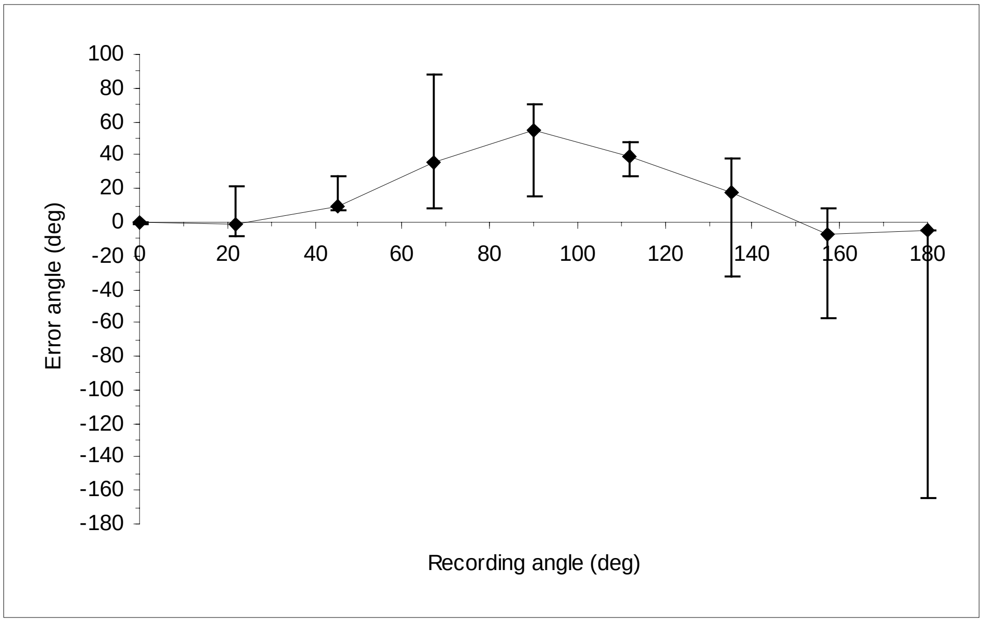


(b) Array with varied psychoacoustics

Figure 12: Test 1: Omni arrays with narrow front coverage angle: Median values and associated 25th to 75th percentile of error angle vs. recording angle

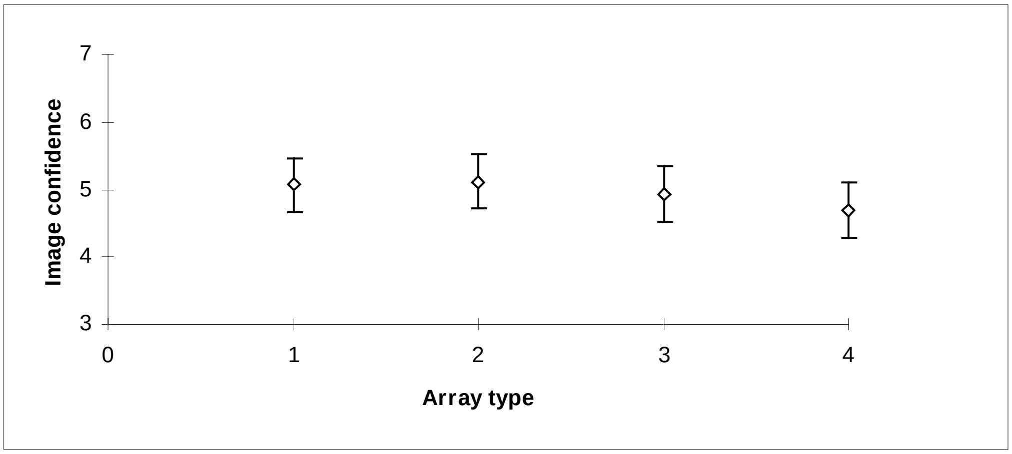


(a) Array with constant psychoacoustics

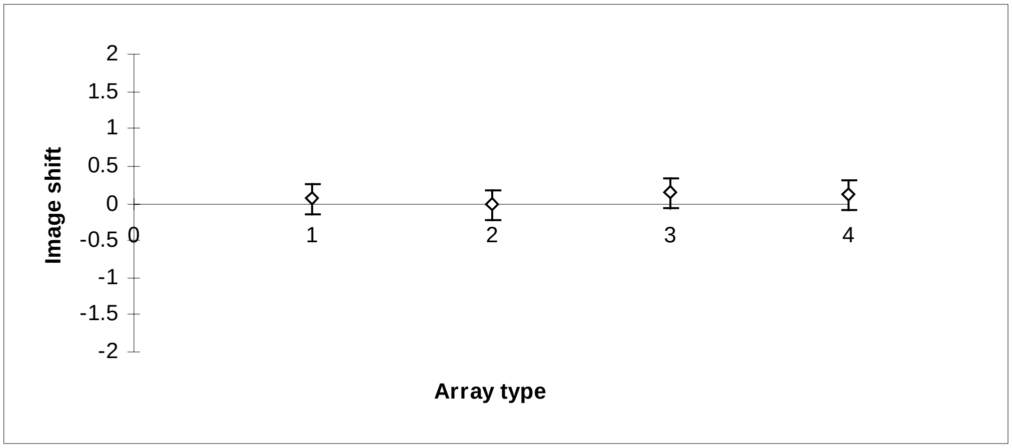


(b) Array with varied psychoacoustics

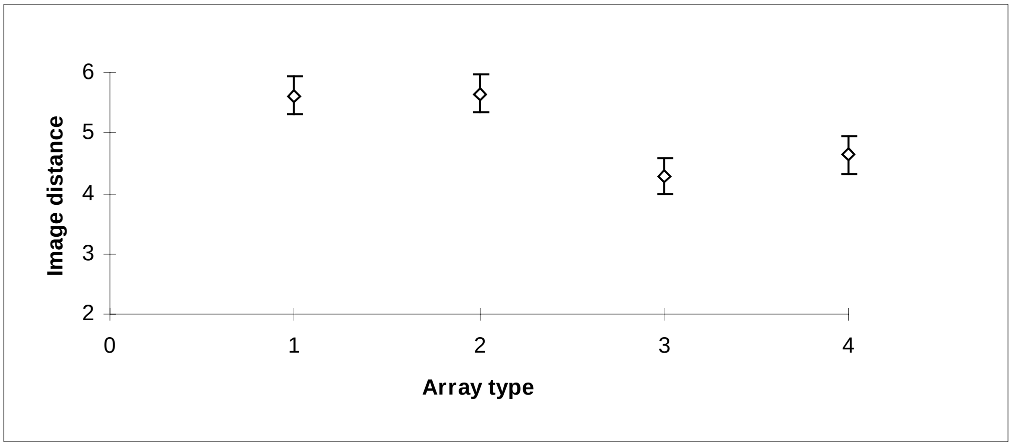
Figure 13: Test 2: Omni arrays with wide front coverage angle: Median values and associated 25th to 75th percentile of error angle vs. recording angle



a) Image confidence



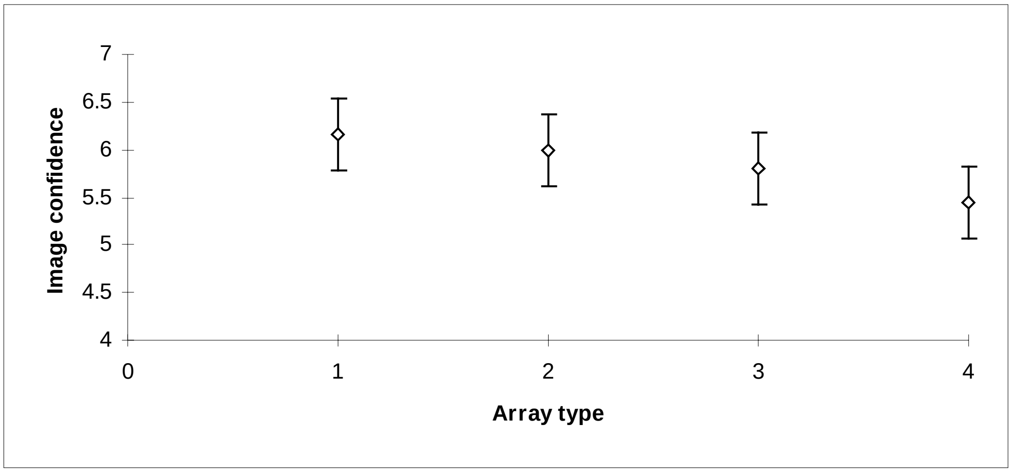
(b) Image shift



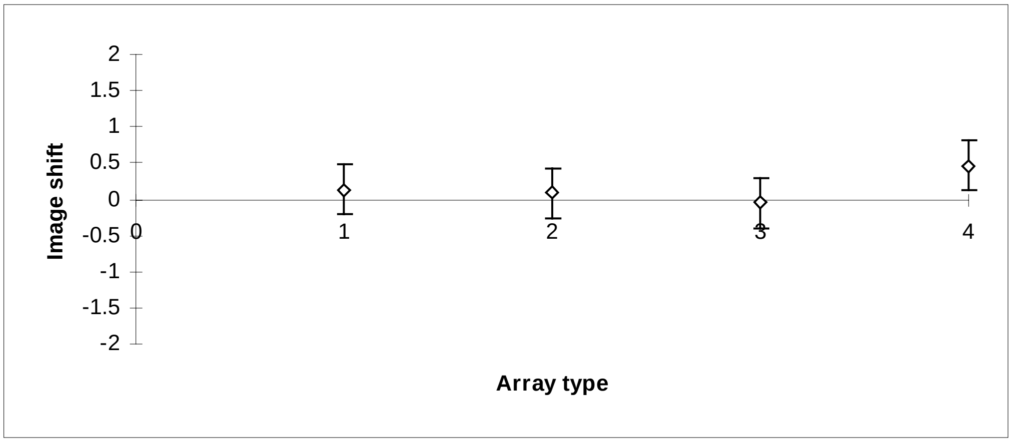
(c) Image distance

Figure 14: Mean values and associated 95% confidence intervals for arrays with narrow front coverage angles as generated by the ANOVA model

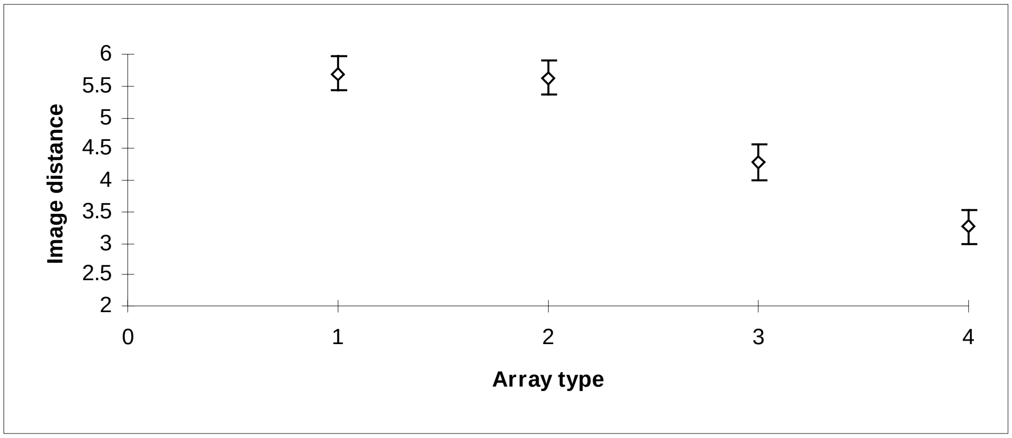
Key: *1 = cardioid with constant psychoacoustics*
2 = cardioid with varied psychoacoustics
3 = omni with constant psychoacoustics
4 = omni with varied psychoacoustics



(a) Image confidence



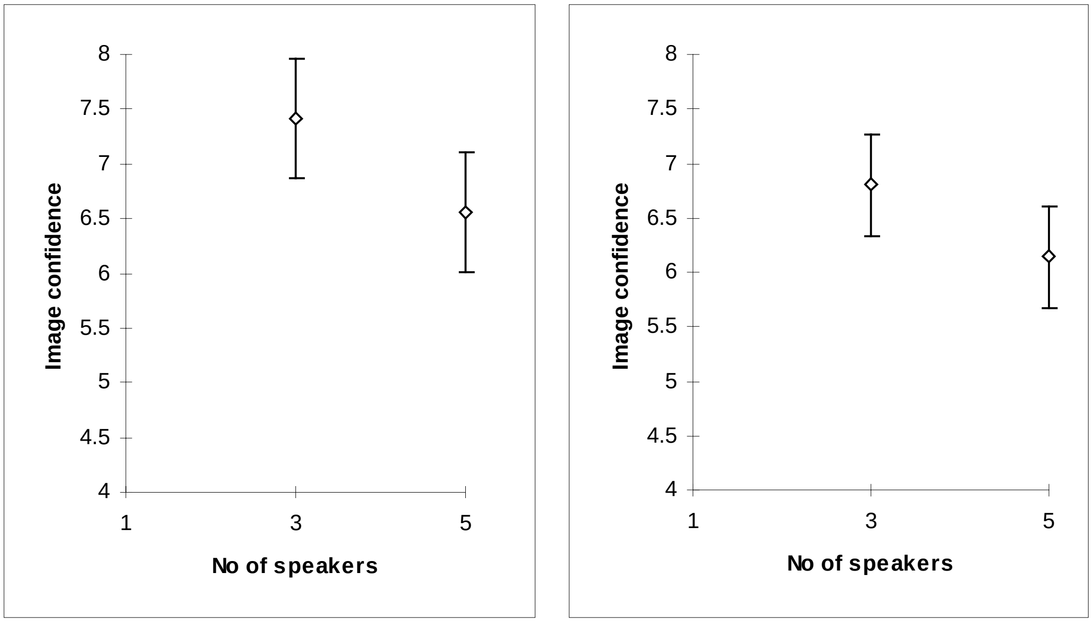
(b) Image shift



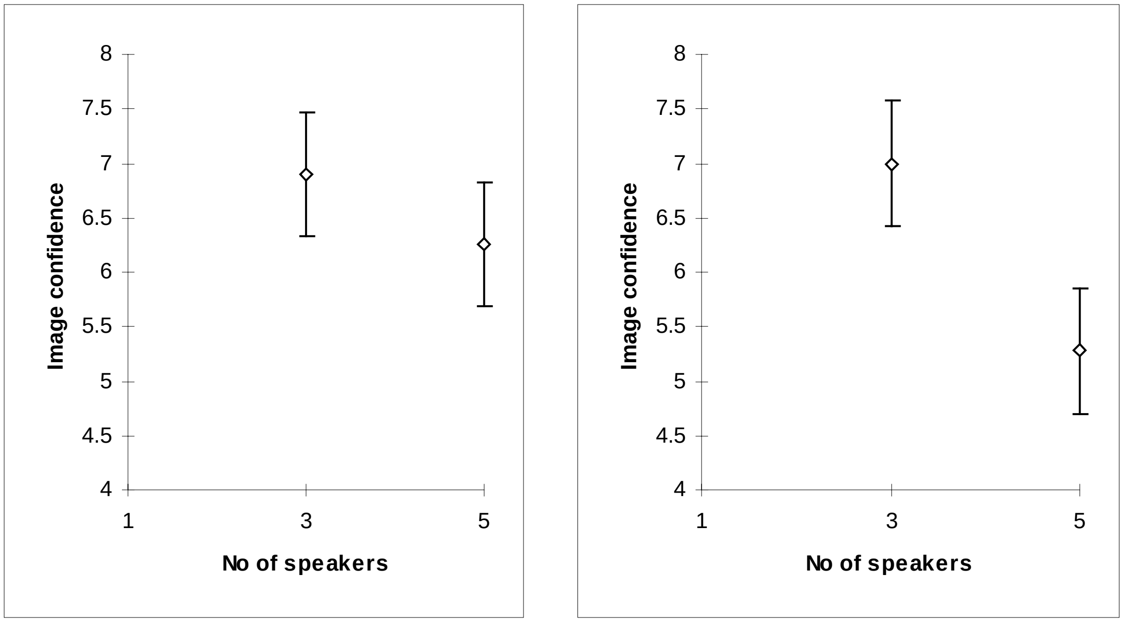
(c) Image distance

Figure 15: Mean values and associated 95% confidence intervals for arrays with wide front coverage angles as generated by the ANOVA model

Key: *1 = cardioid with constant psychoacoustics*
2 = cardioid with varied psychoacoustics
3 = omni with constant psychoacoustics
4 = omni with varied psychoacoustics



(a) Cardioid array with narrow front coverage (b) Cardioid array with wide front coverage



(a) Omni array with narrow front coverage (b) Omni array with wide front coverage

Figure 16: Mean values and associated 95% confidence intervals of image confidence between images produced by 3 and 5 loudspeakers averaged over all subjects for omni arrays as generated by the ANOVA model

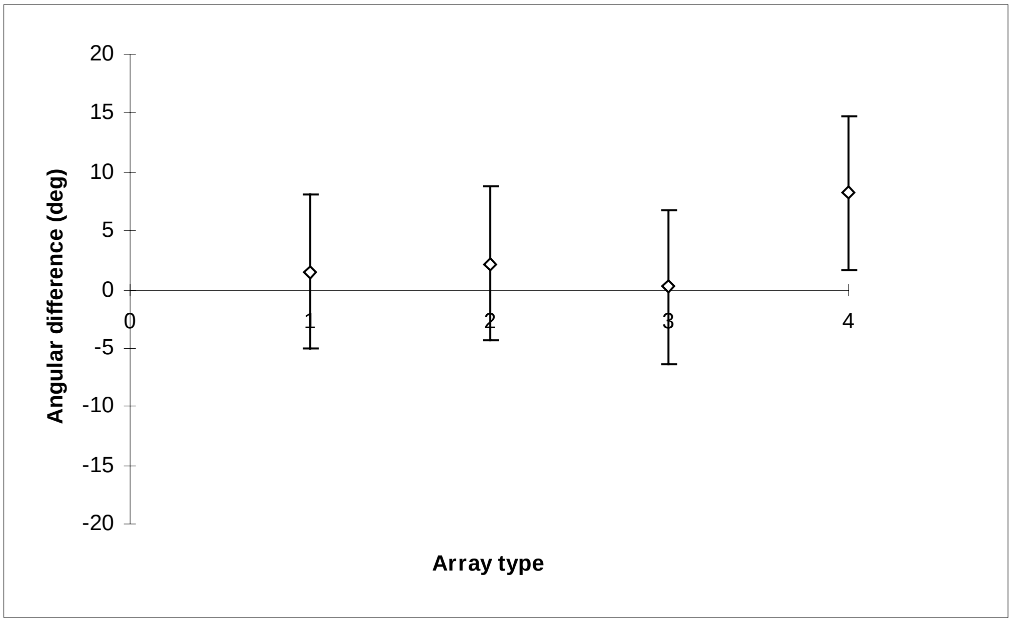


Figure 17: Mean values and associated 95% confidence intervals of angular differences between images produced by 3 and 5 loudspeakers for cardioid arrays as generated by the ANOVA model

Key:

1 = cardioid array with narrow front coverage angle
2 = cardioid array with wide front coverage angle
3 = omni array with narrow front coverage angle
4 = omni array with wide front coverage angle

APPENDIX A. INSTRUCTIONS FOR LISTENERS

Thank you very much for taking part in these listening tests. You will be asked to judge a number of audio extracts, which have been recorded on various designs of surround sound microphone arrays.

The tests themselves are run using the new ALEX software, and are constructed of about 20 identical tests. In each test you will be asked to listen to four audio samples, which you can select at random. The screen will show four buttons, A, B, C, D corresponding to the four samples, and clicking on any of these will start a looped sound. There is also a Pause button to pause the test at any time. When you have rated all the four samples on the screen, move on to the next test by pressing Next at the bottom of the page. NB Please note that you cannot go back to a previous test once you have left it.

The individual audio samples need to be rated against four criteria, using the sliders shown on the screen. These criteria are:

- ***image location*** - to indicate from where in a horizontal plane around you the sound appears to be coming

-180	-90	0	90	180
sound from directly behind the head	sound at right angle to the left	sound from directly in front	sound at right angle to the right	sound from directly behind the head

- ***localisation confidence*** - to rate how confident you are in establishing the image position, whether the image is stable or diffuse

0	2.5	5	7.5	10
no stable image formed	diffuse image, from a general direction	reasonable image, which may move with head movement	good image, reasonably stable with head movement	sharp image, clearly localised and stable

- ***image shift*** - whether the image shifts left or right between the initial sound transient and the sound decay

-5	-2.5	0	2.5	5
image shifts noticeably to the left	image shifts slightly to the left	image does not shift	image shifts slightly to the right	image shifts noticeably to the right

- ***image distance*** - whether the image sounds close or farther away

0	5	10
image seems close	image at middle distance	image seems far away

Entering the ratings for each audio sample is done by clicking on the arrow on the slider, and dragging it to the desired point. This can be anywhere on the scale. It is not the aim to compare the four samples in each test, and none of them constitute a reference.

When listening to the audio samples, you may find that the image can only be localised by moving your head, which you are free to do. However, when deciding on your final scoring, please do so when facing generally forward, and with your head not more than about ±20° from the straight ahead direction. Above all, please be honest about what you hear - it is after all your opinion that we are after - there are no right or wrong answers.