Audio Engineering Society
Convention Paper 5673
Presented at the 113th Convention
2002 October 5–8 Los Angeles, CA, USA

Localization of lateral phantom images in a 5-channel system with and without simulated early reflections

Jason Corey and Wieslaw Woszczyk

Mulitchannel Audio Research Lab, Faculty of Music, McGill University, Montreal, QC, Canada, H3A 1E3

Correspondence should be addressed to Jason Corey (jason.corey@mail.mcgill.ca)

ABSTRACT

Phantom images that rely on interchannel level differences can be produced easily for two-channel stereo. Yet one of the most difficult challenges in production for a five-channel environment is the creation of stable phantom images to the side of the listening position. The addition of simulated early reflection patterns from all five loudspeakers influences the localization of lateral phantom sources. Listening tests were conducted to compare participants' abilities to localize lateral sources under three conditions: power-panned sources alone, sources with simulated early reflection patterns, and simulated early reflection patterns alone (without direct sound). Results compare localization error for the three conditions at different locations and suggest that early reflection patterns alone can be sufficient for source localization.

0 INTRODUCTION

When producing music and film soundtracks for five-channel (3/2) reproduction, there is a possibility to create the impression of sound sources emanating not only from the front but also from the side and behind the listener. Because of the wide aperture between front and rear loudspeakers (80-90°) in a five-channel configuration (ITU-R BS.775) [1], and due to the loudspeakers projecting sound to one side of the listener's head, it is difficult to create a stable lateral sound image using only interchannel level differences.

Simple constant-power pair-wise panning, perhaps the most commonly used panning algorithm, is not as reliable for lateral positions as it is for the front. Localization of phantom images between loudspeakers in front of a listener (summing localization) has been studied by Blauert [2] and Griesinger [3]. Phantom images can be produced in conventional two-channel stereo when the loudspeakers are symmetrically arranged (normally ±30°) across the median plane in front of a listener. Phantom images are perceived between loudspeakers emitting identical signals, equidistant from a listener. By changing the amplitude of one loudspeaker relative to the other, it is possible to change the perceived location of the phantom image. Figure 1 illustrates the signal path from each speaker to the ears of the listener. Normally constant-power (or sine/cosine) panning is used to position phantom images in stereo reproduction systems.

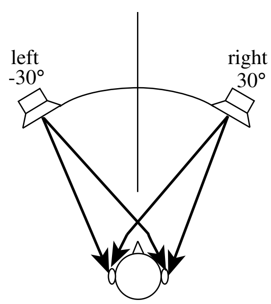

Fig. 1: Signal paths from loudspeakers to ears in a conventional stereo (2/0) system.

Studies in auditory perception have determined that localization of real sound sources is accomplished through interaural time differences (ITD), interaural level differences (ILD), and head-related transfer functions (HRTF's). [4] It was reported by Blauert [2] that there is a large variability in localization of real sources for lateral positions as shown in Figure 2. This may in part be explained by the "cone of confusion," where front-back confusions are made in the localization of lateral sources that are slightly ahead or behind the listener. There is considerable localization blur for sources at ±90° which indicates some variability in source localization and a perception of the source being wider than it is physically. Even before phantom source localization is considered it needs to be emphasized that there is a substantial localization blur even for real sources.

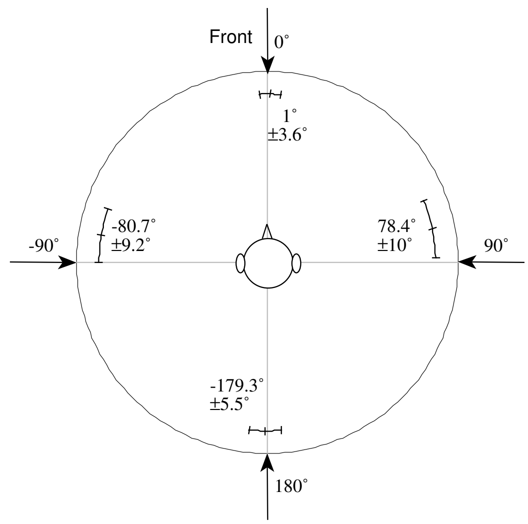

Fig. 2: Localization blur and localization in the horizontal plane for loudspeakers emitting white-noise pulses of 100ms duration at $\phi = 0°, \pm 90°$, and 180°. Arrows indicate the location of the sources. (Diagram adapted from Blauert [2], p. 41, after Preibisch-Effenberger [5] and Haustein & Schirmer [6])

In an investigation of localization of lateral phantom images Theile and Plenge [7] found a large variations in the perceived location of lateral phantom images. Damaske [8] has also indicated that it is generally difficult to create lateral phantom images. No matter what the amplitude difference between the loudspeakers, the ipsilateral ear will always receive the signals from both loudspeakers at a higher amplitude and with less time delay than the contralateral ear. Perhaps the auditory system is not relying on ILD and ITD for localization because there is little change in the interaural level or time difference as a source is power panned from a loudspeaker at 30° to one at 120°. Pinna filtering may play a larger role in this case. Figure 3 illustrates the sound propagation paths from side loudspeakers to the ears of a listener seated in the centre of the array. Whenever pairs of adjacent loudspeakers are not arranged symmetrically across the median plane, it is difficult to rely on interaural differences to determine location.

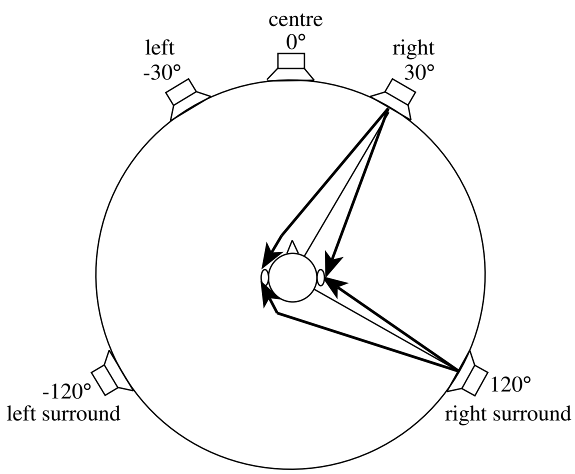

Fig. 3: Signal paths from loudspeakers to ears for a lateral power-panned phantom image in a 3/2 channel system.

Despite the claimed difficulty in generating lateral phantom images, Lund [9] has suggested that the stability and localization certainty of lateral phantom images in a five-channel reproduction environment can be improved by rendering early reflection patterns from all loudspeakers. As this was a preliminary investigation, the current paper presents additional experimental methods and results to help understand the relationship between simulated early reflection patterns and the perception of lateral phantom images.

Rather than suggest the addition of more channels on the side, the challenge here is to work with the existing ITU loudspeaker layout standard and find better methods of signal processing to improve imaging. Because of the difficulty in producing stable lateral phantom images, early reflection patterns need to be generated from all loudspeakers that "support" and stabilize the perceived location of the direct sound. It was deemed that adding early reflection patterns, calculated according to the direct source location, would produce a less ambiguous perceived source location because there will be more information for the auditory system to judge the location of the direct source. This is consistent with Casey [10] who states that there is an inverse relationship between the complexity of a stimulus and its respective perception. The more complex a stimulus is, the easier it is for a listener to draw conclusions about it. With a simple stimulus it is more difficult to determine information about the source through the perception of the physical event. A power-panned source with rendered early reflections will create a much more complex stimulus than the source without reflections, and perhaps will create a more stable perceived source location.

The paper will describe listening tests conducted to further investigate the influence of simulated early reflection patterns from all five loudspeakers on the perceived location of a lateral phantom image. Results will compare the localization errors for the three different simulated acoustic conditions. The research questions for the paper are thus:

- how well does constant-power panning work for the localization of lateral sources?
- how do simulated early reflection patterns (rendered according to source location in a given room model) influence the localization of lateral phantom images in a five-channel environment?
- can lateral localization in a five-channel system be improved by rendering early reflection patterns to support the direct sound?
- does an image model provide the best location support for the direct sound?
- was the image model chosen, the correct one?
- can listeners localize sounds in a five-channel environment with simulated early reflection patterns alone (no direct sound)?
- is listener localization accuracy related to certainty about the image location?

1 DESCRIPTION OF THE LISTENING TEST

1.1 The Listening Room

Listening tests were conducted in an acoustically damped room with a 3/2 loudspeaker layout according to the ITU-R BS.775 recommendation [1]. Five two-way active loudspeakers (Bang & Olufsen Beolab 4000) were placed at 0°, ±30°, and ±120° with a radius of 2.1 m from the central listening position. An acoustically transparent, visually opaque curtain was hung directly in front of the loudspeaker array. It served to conceal the location of the loudspeakers from the listeners so that the participants' auditory perceptions would not be influenced by the visual perception of speaker positions. It is well known that when making judgments about sound source location that the auditory system can be influenced quite strongly by visual judgments of source location. A good literature review of auditory-visual interaction can be found in [11].

The ambient sound level in the room with the computer running was 27dB SPL, A-weighted, fast time weighting, measured from central listening position. Calibration of the loudspeaker levels was performed using a Brüel & Kjær 2235 sound level meter, with A-weighting and fast time weighting. Distance of the loudspeakers from the central listening position was calibrated by comparing time delay measurements using a MLSSA. The room reverberation time is approximately 0.270 sec at 125 Hz, 0.172 sec at 250 Hz and 0.168 sec at 500 Hz as noted in [12].

Participants were seated behind a desktop computer display in the centre of the loudspeaker array. The computer was used to automate the test, record the participants responses, and perform the panning and generation of early reflection patterns in real-time. The CPU of the computer was placed in an acoustically isolated machine room to minimize the noise floor of the room.

1.2 Participants and Test Duration

Fifteen subjects, all students and faculty lecturers in sound recording at McGill, participated in the experiment. The test duration was around 30–45 minutes, and all participants were remunerated for their time. All participants completed the test twice at two different times. For this reason, it is considered a repeated measures experiment. During the test, subjects could mute the sound at any time to take a short break if needed. Each subject went through a few examples before beginning the test to familiarize themselves with the procedure and to have a chance to ask questions for clarification of the task. Participants had normal hearing although it was not tested.

1.3 Method

Although other panning algorithms exist such as vector-based amplitude panning [13], Ambisonics [14], and polarity-restricted cosine [15], the most commonly used panning algorithm, constant-power (sine/cosine), was chosen to position the direct sound. Sources were positioned to locations to the side of the listening position, i.e., between left and left-surround and between right and right-surround loudspeakers. Gains (g) applied to the front and rear loudspeaker signals were derived from Equations 1 and 2 where ϕ = intended angle (from 30°–120°):

$$g_{front} = cos(\phi - (\pi/6)) \tag{1}$$

$$g_{rear} = sin(\phi - (\pi/6)) \tag{2}$$

Early reflection patterns were generated using a two-dimensional (horizontal plane) image model [16] for each source location and panned to the loudspeakers using the same sine/cosine panning algorithm. Although the direct sound always originated from at most two loudspeakers, the early reflections were rendered through all five loudspeakers. Early reflections, generated in real-time using a desktop computer, were calculated up to 4th-order in the horizontal plane only, giving 40 reflections in total. Reflections were low-pass filtered to approximate wall and air absorption, with each successive order of reflections having a lower cut-off frequency. Specifically the cut-off frequencies were:

- 1st order reflections: $f_c = 16$ kHz
- 2nd order reflections: $f_c = 12$ kHz
- 3rd order reflections: $f_c = 10$ kHz
- 4th order reflections: $f_c = 8$ kHz

The delay time and gain according to signal propagation for each reflection was calculated according to the distance travelled, where:

$$delay = distance/c;\ c = 344m/s \tag{3}$$

$$gain = 1/distance \tag{4}$$

Participants were given a graphical user interface on a computer display representing a top view of the room (Figure 4) and were asked to position a black dot to a location on the interface that best represents the location of the sound source. This is nearly identical to the method used by Martin *et al.* [17] for testing perceived location of phantom images, where listeners were asked to place a dot on a graphical interface to indicate the perceived location of a phantom image.

Preliminary versions of the interface included dots to indicate ±30°, ±60°, ±120° with corresponding markers placed around the room to give listeners points of reference. It was decided that these dots and markers would influence the listeners by introducing some unwanted, artificial quantization in the responses and it might also give some information away about what was being tested. Markers around the room would require the subjects to turn around for a visual indication of

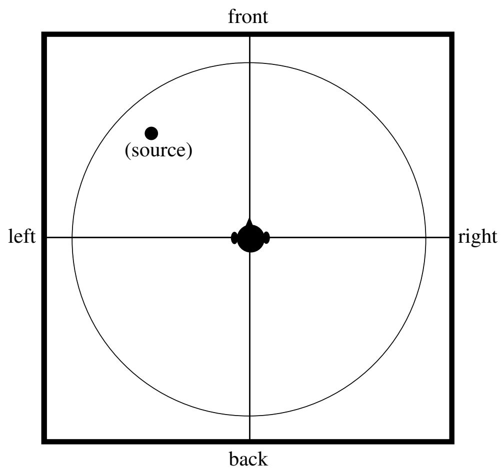

Fig. 4: Graphical interface in which participants indicated, with a black dot, the perceived location of the sound image.

the marker locations, thus making it difficult to make an accurate decision about their auditory perception.

As Evans states [18], having listeners place a dot to indicate source location has the advantage of not requiring the listener to translate a direction into a verbal response. The method of having a graphical interface with which to indicate source location may be better suited than having a pointer because participants can indicate rear locations without turning around to look. This is especially important because the method of generating lateral phantom images is sensitive to head orientation.

Participants heads were not fixed, but they were instructed to indicate perceived source location while facing forwards. A marker was placed at 0° front centre on the ceiling near the acoustically transparent curtain to indicate this location in the room. By facing the computer screen directly, they could be sure that they were facing forwards. They were shown a mark on the ceiling above their head indicating the centre of the room.

The reproduction level was approximately 65 dB SPL, linear frequency weighting, slow time weighting.

1.4 Independent variables

For the test three monophonic, anechoic sound sources were used:

1. speech, female Danish
2. percussion, bongos
3. electric guitar

The speech and percussion samples were taken from the Bang & Olufsen *Music for Archimedes* compact disc [19], which were recorded in an anechoic chamber. The guitar sample was recorded using a microphone placed very close to the guitar amplifier, providing isolation from the room. The choice of sound sources represents the fulfillment of three criteria: a transient source (percussion), a more steady-state source (electric guitar), and a third source that exhibits a mixture of transient and steady-state characteristics (speech).

Three "room effect" conditions were presented:

1. dry, anechoic source, positioned using constant-power panning (sin/cos)
2. dry sound with early reflections (4th order image model, horizontal plane only)
3. early reflection pattern only

Eight source locations were tested (Figure 5):

1. −45° (left)
2. −65°
3. −80°
4. −100°
5. 45° (right)
6. 65°
7. 80°
8. 100°

It was decided to treat the left and right sides separately and not to average the two together. In this way, the test results might illustrate differences in responses for left and right, indicating asymmetrical room effects in the listening room and/or differences in perception of the two sides.

The intended angle 45° was chosen to represent a location just outside of the front loudspeaker at 30°. As it was known by the authors from initial experimenting and listening that this position tended to pull towards the closest loudspeaker, it was decided that this would be a crucial location to test.

The intended angle 65° was chosen to determine if listeners could differentiate between 45° and a more lateral location. Also it was found through informal listening before the test that this location was slightly peculiar in that the sound source appeared to pull out from the loudspeakers towards the listening position. It was thought

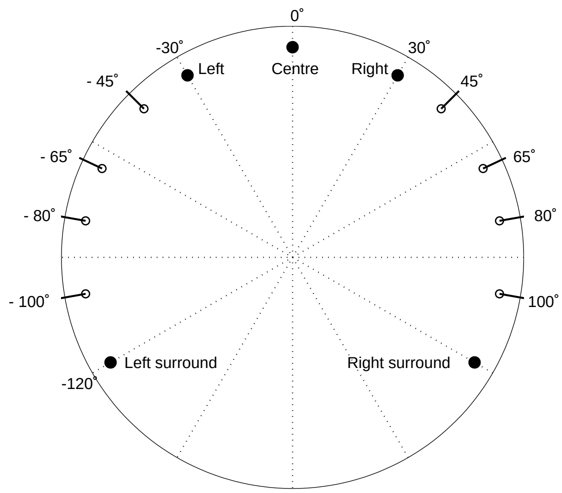

Fig. 5: The eight lateral source locations tested, relative to the five speaker locations (indicated by black dots).

that this might influence the perception of azimuth differently for the "room effect" conditions.

Finally, to test listeners' abilities to differentiate between locations slightly ahead and behind 90°, positions at 80° and 100° were chosen. We wanted to find out if these locations would be quantized to the same lateral auditory event. Or if an intended angle of 100° would be perceived to pull back towards the rear loudspeaker.

Three sound sources, three room effect conditions, and eight locations makes a total of 72 permutations. The listening test was fully factorial for each participant, with random order presentation of the variable combinations. The test was repeated once for each listener.

To provide additional information about the listeners indication of source location, they were also asked to rate their "certainty" of the source location on a five-point scale, where 5 is most certain, and 1 is least certain. Specifically they could choose one of the five following levels of certainty:

- 5 – no doubt
- 4 – high
- 3 – good
- 2 – some
- 1 – poor

This is similar to Lund's [9] listening test design where listeners were asked to rate certainty of source location on a five point scale. There they also rated "robustness" and "diffusion" to determine a composite "consistency score". By using the certainty rating scale in addition to asking for a specific indication of perceived location, it is possible to compare certainty with accuracy. It was hypothesized that as certainty ratings increase, accuracy of localization would also increase.

Listeners response times to place the dot were also recorded. This was done without their knowledge so that it would not influence their responses, or make them feel rushed. It was deemed that response times should correlate negatively with certainty ratings, and that it might provide additional information about the difference between localization of direct sound and direct with reflections.

2 RESULTS

An 8(intended angle) * 3(room effect) * 3(sound source) repeated measures analysis of variance of within-subjects effects was performed with the results presented in Table 1.

The following independent variables and interactions had significant effects on the dependent variable perceived angle (with respective F- and p-values found in Table 1):

- intended angle
- sound source
- intended angle * room effect interaction
- intended angle * sound source interaction

The following independent variables and interactions had significant effects on the dependent variable response time (with respective F- and p-values found in Table 1):

- intended angle
- intended angle * room effect interaction
- intended angle * sound source interaction
- intended angle * room effect * sound source interaction

The following independent variables and interactions had significant effects on the dependent variable certainty (with respective F- and p-values found in Table 1):

- intended angle

- room effect
- intended angle * room effect interaction
- intended angle * sound source interaction
- intended angle * room effect * sound source interaction

Figures 9–20 illustrate the perceived locations (with 95% confidence intervals) of the sound sources relative to the intended angle and loudspeaker positions, plotted according to sound source and room effect condition. To make the results easier to read only four intended locations (not all eight) are plotted on each figure, alternating between a plot with ±45° and ±80° and a plot with ±65° and ±100°. Figures 9 and 10 indicate the perceived locations of all three sound sources for the anechoic condition. Figures 11 and 12 indicate the perceived locations of all three sound sources for the anechoic with early reflections condition. Figures 13 and 14 indicate the perceived locations of all three sound sources for the condition of early reflection patterns only.

Figures 15 and 16 indicate the perceived locations of the speech source for all three room effect conditions: anechoic, anechoic with reflections, and reflections only. Figures 17 and 18 indicate the perceived locations of the percussion source for all three room effect conditions. Figures 19 and 20 indicate the perceived locations of the electric guitar source for all three room effect conditions.

Figures 21–23 show the mean certainty ratings (with 95% confidence intervals) for each of the three room effect conditions, plotted according to intended angle.

Figures 24–26 show the mean response time (with 95% confidence intervals) for the three room effect conditions, plotted as a function of intended angle.

3 DISCUSSION

As the results indicate the independent variable intended angle had significant effects on the perceived angle. It is expected that listeners should perceive the eight distinct intended locations. Intended angle also had significant effects on certainty and response time. Figures 21–26 illustrate the differences in response time and certainty as a function of intended location for the three room effect conditions.

We would expect there to be a negative correlation between response time and certainty of source location which there was. As was consistent with Lund [9], the room effect condition had a significant effect on certainty.

The localization blur of intended locations ±65°, ±80°, and ±100° is quite close to what Blauert [2] reported for real sources (see Table 2). Often the localization blur is similar for the anechoic and anechoic with reflections conditions.

It was hypothesized that by adding simulated early reflection patterns to the constant-power panned anechoic sound, that this would decrease the standard deviation of the perceived location of the phantom image. Generally it does not seem to be the case although there are exceptions such as for speech at ±80°, ±65° and 100° right in Figures 15 and 16, for electric guitar at 100° right in Figure 20. The question is: why not? Some possible explanations are as follows.

The production of lateral phantom images using constant-power panning is much more difficult than it is in a traditional two-channel stereo setup due to the fact that the signals from both speakers are reaching the ipsilateral ear with a shorter time delay and with greater intensity that the contralateral ear (Figure 3). As such, when a source is panned between the side loudspeakers, the image is not necessarily perceived as having the same "width" and different frequency bands may be localized differently. For instance it has been found empirically by the authors (and also reported by some participants) that for a single source panned to a lateral position (e.g., 90°), the high frequency band (typically above 1500 Hz) is localized near the front loudspeaker and the low frequency band is localized towards the side or rear. As such the exact location of the phantom image can be ambiguous. Participants involved in the listening test were required to place a single dot to indicate the perceived location of the phantom image. They may have had trouble deciding how to represent a wide phantom image with a small dot. Perhaps the relative balance of the high and low frequency bands influenced their decision.

The results also indicate evidence of what Gerzon refers to as the "detent" effect [14] where sound images that are intended to be near a loudspeaker but not directly at a loudspeaker (e.g., left 45°) are pulled towards that loudspeaker (left 30°). This is apparent in Figures 9, 11, and 13 for all sound sources and room effect conditions. The speech sound source, for all room effect conditions, is generally perceived to be closest to the intended angle of ±45° of the three sounds. In addition it may be noted that the ±45° intended locations resulted in smaller standard deviations of the perceived locations for all sound sources and all room effect conditions, as compared to other intended locations. The certainty ratings are higher (Figures 21–23) and the response times are generally shorter (Figures 24–26) for the intended location of ±45°. The intended angle 65° also exhibited the "detent" effect as perceived locations were pulled towards the loudspeakers at ±30°.

From the results it is apparent that different types of sources are localized differently. This is consistent with the findings of Blauert [2] and Griesinger [3] who found that the perceived position of a phantom image is strongly dependent on the spectrum of the source. For instance the electric guitar source was in many cases (e.g., Figures 9 – 14) localized more towards the front loudspeaker than the other two sound sources. This may be explained by the fact that it has a higher spectral centroid than both the percussion and speech samples. Due to the nature of pinna filtering, we are more sensitive to high frequencies located in front. Figure 6 illustrates the magnitude spectrum of the electric guitar sound sample, calculated using a 1024-point hanning window DFT averaged over the entire sound file. Comparing the magnitude spectrum of the speech (Figure 7) and percussion (Figure 8) samples, it is possible to see that there is considerably less energy (approximately −30dB) in the frequency band from 2–3 kHz, than there is in the electric guitar sample. The temporal characteristics of the sound sources may also have influenced the judgment of location.

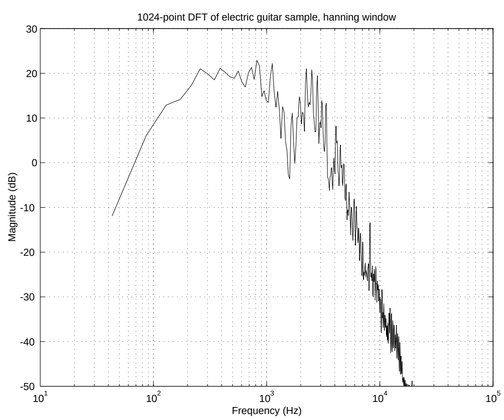

Fig. 6: Magnitude spectrum of the electric guitar sample, averaged over the length of the sample.

As can be seen from Figures 13 and 14, participants were generally able to localize the sound sources with the early reflection patterns only (no direct sound). In this case the electric guitar is not as accurately localized as the speech and percussion.

In comparing the localization accuracy with certainty, it was found that there was not a significant correlation between the variables. From this we can conclude that

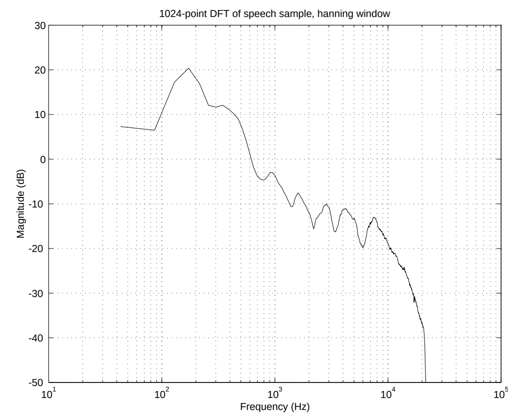

Fig. 7: Magnitude spectrum of the speech sample, averaged over the length of the sample.

confidence in source location does not always translate into accurate or consistent localization ability.

Some subjects reported afterwards that a few examples were perceived as having increased elevation over others. Although it is not known which conditions elicited this perception, it would be an interesting study to pursue.

A couple of participants indicated that the visual modality may have influenced their location judgments. It may have been difficult for listeners to translate from the horizontal auditory plane surrounding them to the visual interface which was in front of them and angled nearly vertical. Perhaps markers around the room and corresponding markers on the visual interface would have helped listeners make more accurate and consistent responses. There may also have been some confusion about where true $\pm 90^\circ$ was. One participant indicated that there was a tendency to rely on peripheral vision to determine $\pm 90^\circ$, and this may have skewed the judgments slightly forwards for the intended angle of $\pm 80^\circ$.

One participant noted after the test that in a couple of the examples the sound source was perceived to be in close proximity to the centre of the room, pulling out from the opaque curtain. Although it is not known for certain which conditions elicited this perception, we would guess that it would be the condition with a dry sound source panned to + or -65°.

As with Blauert [2], it should be noted here that the results cannot confirm whether the deviations in localization result in errors in participants judgments about

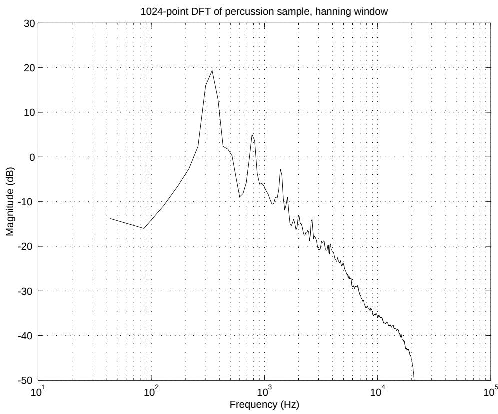

Fig. 8: Magnitude spectrum of the percussion sample, averaged over the length of the sample.

direction or whether they accurately reflect their auditory localization.

4 CONCLUSIONS AND FUTURE WORK

The results of the investigation indicate that the addition of reflections to the direct source had some but not a significant influence on the localization blur of the source, but it did have a significant effect on the certainty rating. Sound source type had significant effects on the perceived location, underlining the importance of using different program material for localization tests. Early reflection patterns alone rendered according to intended source location result in fairly accurate localization of sources.

One possible way to increase the influence of reflections on localization of a dry source might include diffusion of the reflections. Since the room boundaries are treated essentially as mirror reflectors (with some attenuation due to distance and filtering) in this model, simulated diffusion might create more realistic sounding reflections.

In addition to this a more complex room geometry might also prove effective, although it becomes more processor intensive as the room model becomes more complex. The image model was arbitrarily chosen to model rectangular room. It is possible that a more suitable room model exists that would provide reflections that are more "supportive" of direct source localization. Alternatively a perceptually based reflection pattern (i.e., one that is not based on the physical dimensions of a real room) might prove to have benefits as well. As Lund [9] points out, appropriate digital signal processing may enable the positioning of sources within a simulated 5-channel sound field that is more robust perceptually than multichannel microphone techniques in a real room.

The constant-power panning algorithm might also be hindering the localization ability. There may be a more appropriate method of panning the direct sound and reflections. Perhaps there needs to be a different panner for the direct than for the reflections.

Future listening tests might include testing the perceived image width and extent, for the three different room effect conditions. An alternate method of testing localization might be to ask listeners to position the sound source at specific locations as indicated by a graphical interface. In this way they would have control over the panner and could change the perceived location of the sound. It is not known whether this would result in different responses than those found here. There was also the issue raised of perceived elevation of the source for some examples. More work needs to be conducted to investigate the perception of source elevation from loudspeaker arrays with an elevation of 0°.

ACKNOWLEDGMENTS

The authors wish to acknowledge the generous support of the National Sciences and Engineering Research Council of Canada (NSERC); Kim Rishøj, Morten Lave, Thomas Lund and T.C. Electronic; Prof. Søren Bech, Poul Præstgaard and Bang & Olufsen; Prof. Daniel Levitin; and to all who participated in the experiment.

REFERENCES

[1] ITU-R. Multichannel stereophonic sound system with and without accompanying picture. Recommendation BS.775-1, International Telecommunication Union Radiocommunication Assembly, 1994.

[2] J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, Mass., revised edition, 1997.

[3] D. Griesinger. Stereo and surround panning in practice. In *112th Convention of the Audio Engineering Society*, Munich, 2002.

[4] B. C. J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, San Diego, Calif., 4th edition, 1997.

[5] R. Preibisch-Effenberger. *Die Schallokalisationsfähigkeit des Menschen und ihre audiometrische Verwendung zur klinischen Diagnostik [The human faculty of sound localization and its audiometric application to clinical diagnostics].*

dissertation, Technische Universität, Dresden, 1966.

[6] B. G. Haustein and W. Schirmer. Messeinrichtung zur Untersuchung des Richtungslokalisationsvernögens [A measuring apparatus for the investigation of the faculty of directional localization]. *Hochfrequenztech. u. Electroakustik*, 79:96–101, 1970.

[7] G. Theile and G. Plenge. Localization of lateral phantom sources. *Journal of the Audio Engineering Society*, 25(4), 1977.

[8] P. Damaske and Y. Ando. Interaural crosscorrelation for multichannel loudspeaker reproduction. *Acustica*, 27:232–238, 1972.

[9] T. Lund. Enhanced localization in 5.1 production. In *109th Convention of the Audio Engineering Society, Preprint 5243*, Los Angeles, 2000.

[10] M. A. Casey. *Auditory Group Theory with Applications to Statistical Basis Methods for Structured Audio*. Doctoral dissertation, MIT, 1998.

[11] R. L. Storms. *Auditory-Visual Cross-Modal Perception Phenomena*. Doctoral dissertation, Naval Postgraduate School, Monterey, California, September 1998.

[12] G. Martin. *A Hybrid Model for Simulating Diffused First Reflections in Two-dimensional Synthetic Acoustic Environments*. Doctoral dissertation, McGill University, Montreal, Canada, 2001.

[13] V. Pulkki. *Spatial Sound Generation and Perception by Amplitude Panning Techniques*. Doctoral dissertation, Helsinki University of Technology, Helsinki, Finland, August 2001.

[14] M. A. Gerzon. Panpot laws for multispeaker stereo. In *92nd Convention of the Audio Engineering Society, Preprint 3309*, Vienna, 1992.

[15] G. Martin, W. Woszczyk, J. Corey, and R. Quesnel. Controlling phantom image focus in a multichannel reproduction system. In *107th Convention of the Audio Engineering Society, Preprint 4996*, New York, 1999.

[16] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

[17] G. Martin, W. Woszczyk, J. Corey, and R. Quesnel. Sound source localization in a five-channel surround sound reproduction system. In *107th Convention of the Audio Engineering Society, Preprint 4994*, New York, 1999.

[18] M. J. Evans. Obtaining accurate responses in directional listening tests. In *104th Conference of the Audio Engineering Society, Preprint 4730*, Amsterdam, 1998.

[19] Bang and Olufsen. *Music for Archimedes*. CD B&O 101, 1992.

Source	Dependent Variable	Wilk's Lambda	F	Hypothesis dF	Error dF	Sig.
A: intended angle	perceived angle	.017	93.943	14.000	194.000	$< .001$
	response time	.441	7.002	14.000	194.000	$< .001$
	certainty	.313	10.906	14.000	194.000	$< .001$
B: room effect	perceived angle	.964	.250	4.000	54.000	.909
	response time	.815	1.456	4.000	54.000	.228
	certainty	.710	2.524	4.000	54.000	.050
C: sound source	perceived angle	.661	3.107	4.000	54.000	.023
	response time	.791	1.675	4.000	54.000	.169
	certainty	.936	.455	4.000	54.000	.768
A * B	perceived angle	.757	2.079	28.000	390.000	.001
	response time	.755	2.097	28.000	390.000	.001
	certainty	.747	2.191	28.000	390.000	.001
A * C	perceived angle	.531	5.186	28.000	390.000	$< .001$
	response time	.804	1.609	28.000	390.000	.028
	certainty	.788	1.766	28.000	390.000	.011
B * C	perceived angle	.873	.962	8.000	110.000	.469
	response time	.832	1.325	8.000	110.000	.238
	certainty	.940	.433	8.000	110.000	.899
A * B * C	perceived angle	.865	1.048	56.000	782.000	.382
	response time	.813	1.519	56.000	782.000	.010
	certainty	.825	1.410	56.000	782.000	.029

Table 1: The results of an 8(intended angle) * 3(room effect) * 3(sound source) repeated measures analysis of variance. Tests of within-subjects effects dependent variables: perceived angle, response time, and certainty.

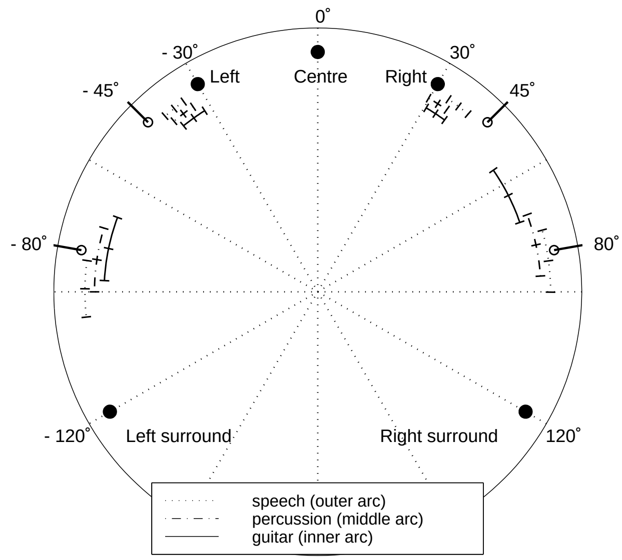

Fig. 9: Plot of mean perceived locations of sound images (with 95% conf. int.) for anechoic condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

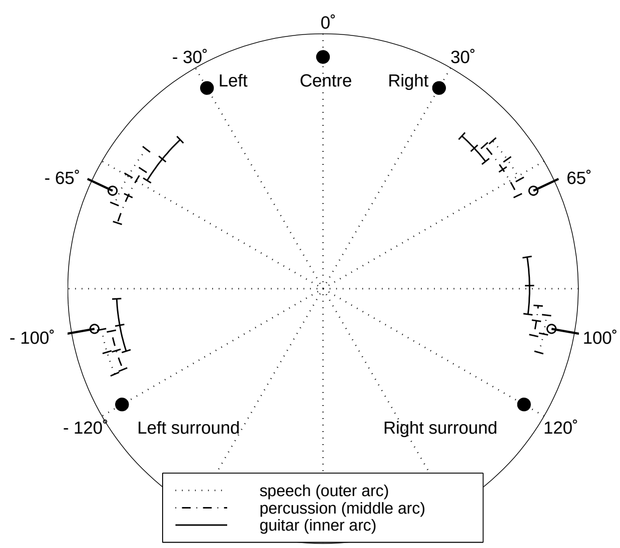

Fig. 10: Plot of mean perceived locations of sound images (with 95% conf. int.) for anechoic condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

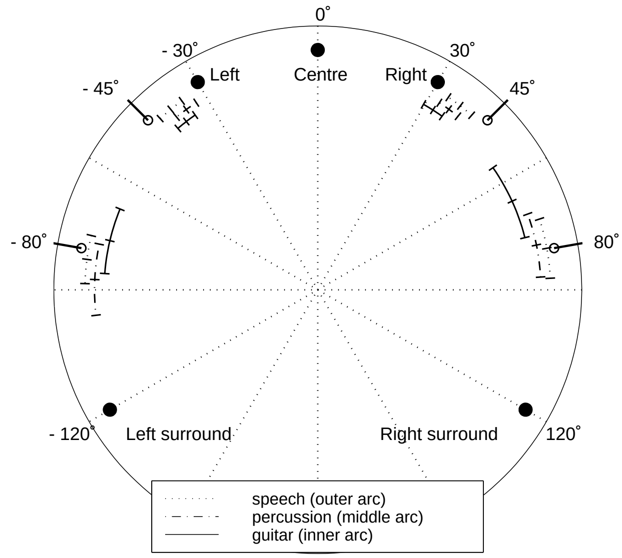

Fig. 11: Plot of mean perceived locations of sound images (with 95% conf. int.) for anechoic with reflections condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

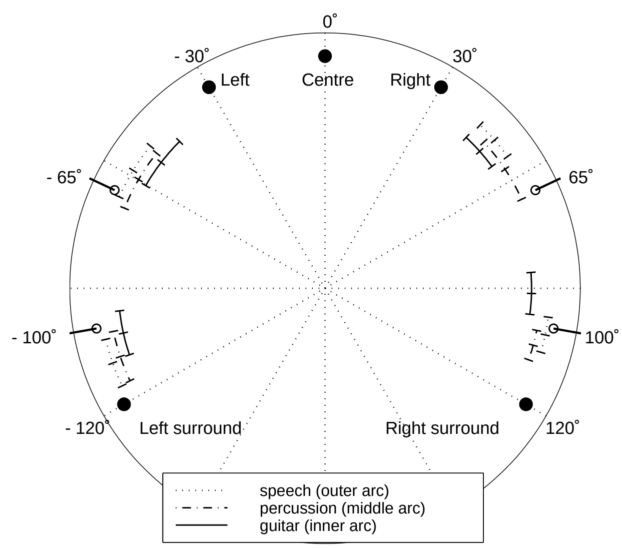

Fig. 12: Plot of mean perceived locations of sound images (with 95% conf. int.) for anechoic with reflections condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

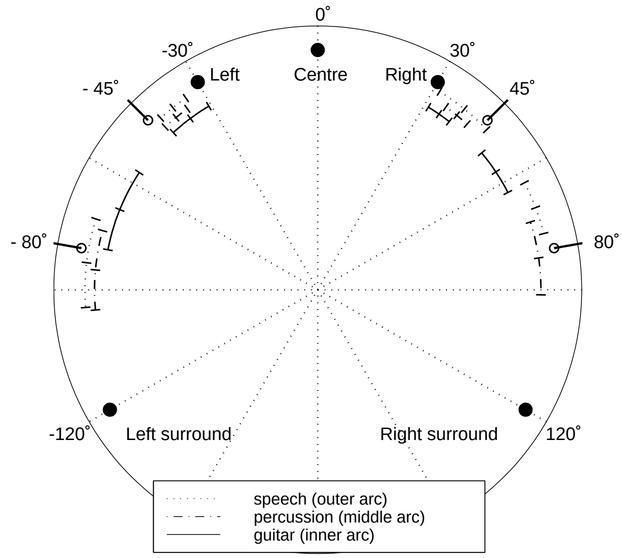

Fig. 13: Plot of mean perceived locations of sound images (with 95% conf. int.) for reflections only condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

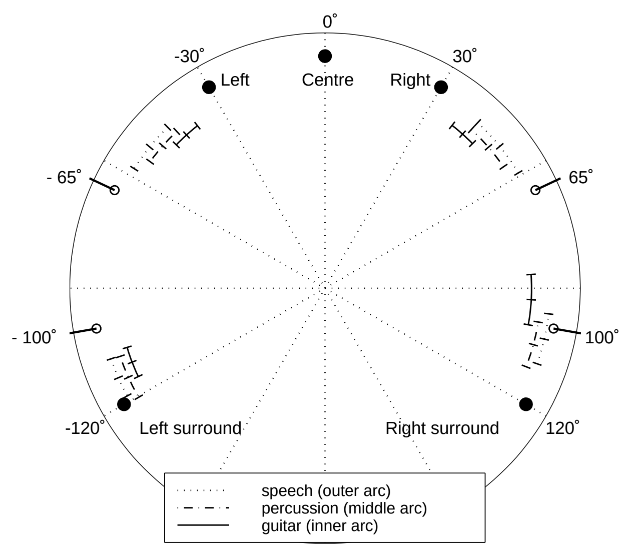

Fig. 14: Plot of mean perceived locations of sound images (with 95% conf. int.) for reflections only condition and all sounds. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

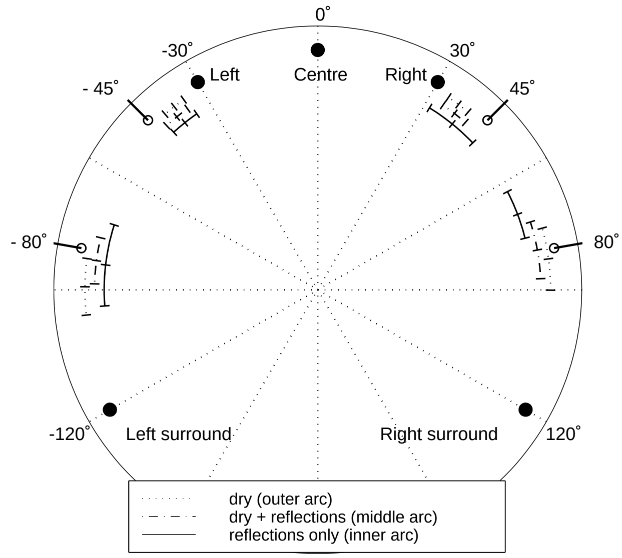

Fig. 15: Plot of mean perceived locations of sound images (with 95% conf. int.) for speech and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

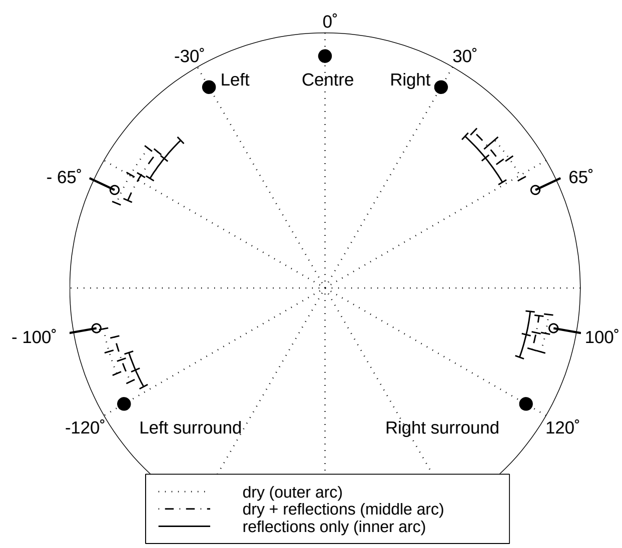

Fig. 16: Plot of mean perceived locations of sound images (with 95% conf. int.) for speech and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

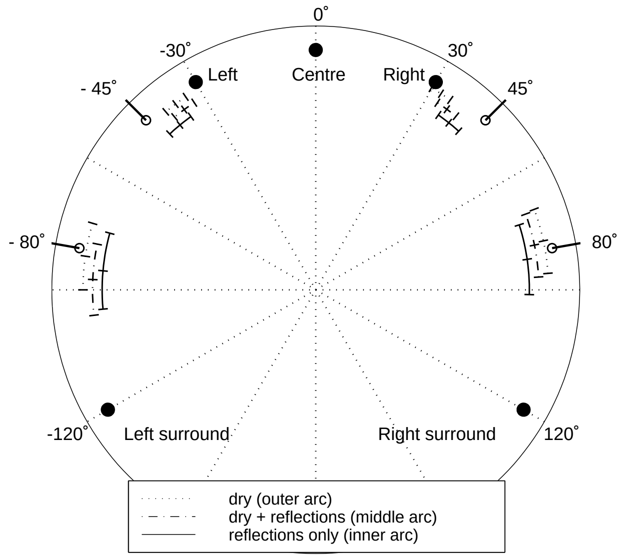

Fig. 17: Plot of mean perceived locations of sound images (with 95% conf. int.) for percussion and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

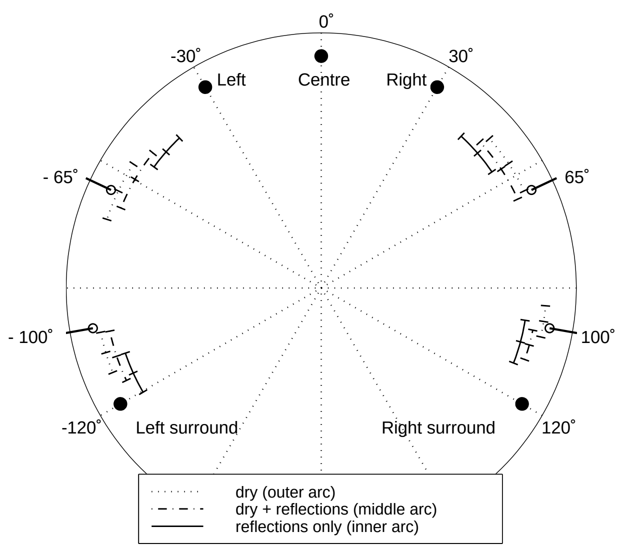

Fig. 18: Plot of mean perceived locations of sound images (with 95% conf. int.) for percussion and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

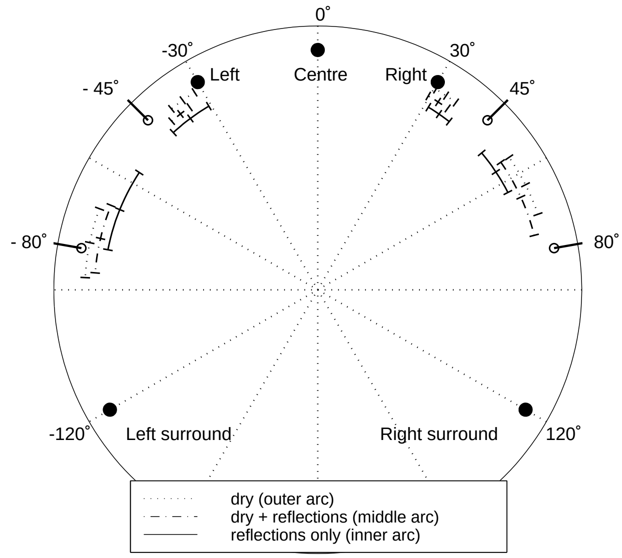

Fig. 19: Plot of mean perceived locations of sound images (with 95% conf. int.) for electric guitar and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±45° and ±80°.

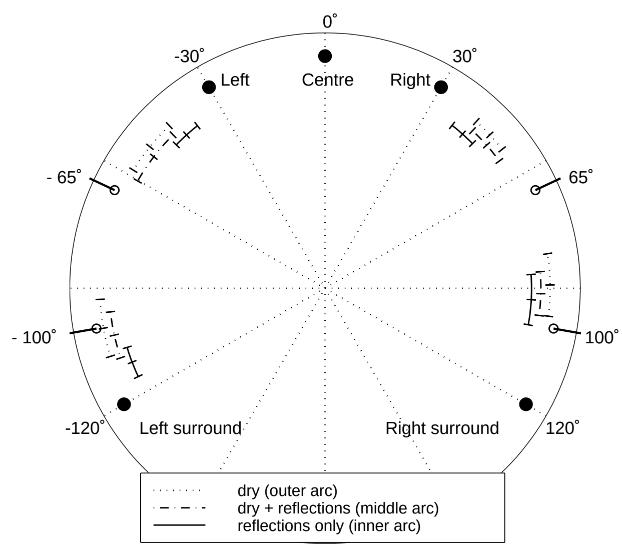

Fig. 20: Plot of mean perceived locations of sound images (with 95% conf. int.) for electric guitar and all room effect conditions. Black dots indicate the loudspeaker locations. Intended image locations were ±65° and ±100°.

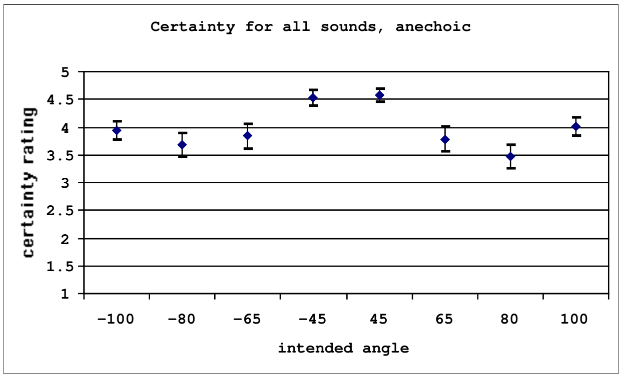

Fig. 21: Plot of mean certainty rating (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the anechoic sound condition.

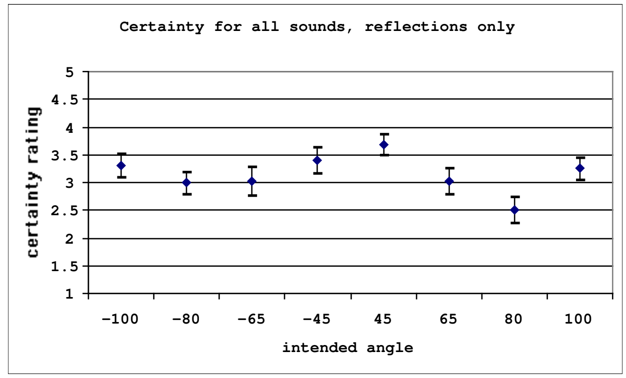

Fig. 23: Plot of mean certainty rating (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the reflections only condition.

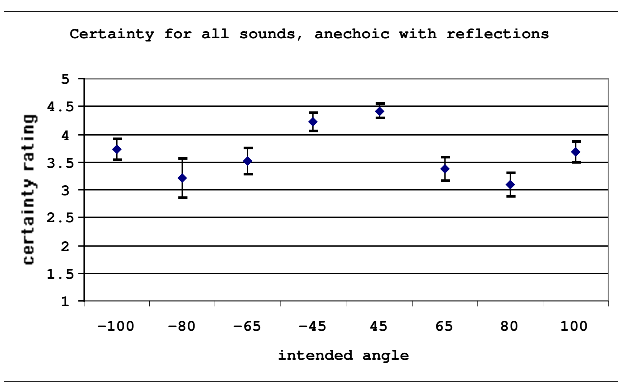

Fig. 22: Plot of mean certainty rating (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the anechoic sound with reflections condition.

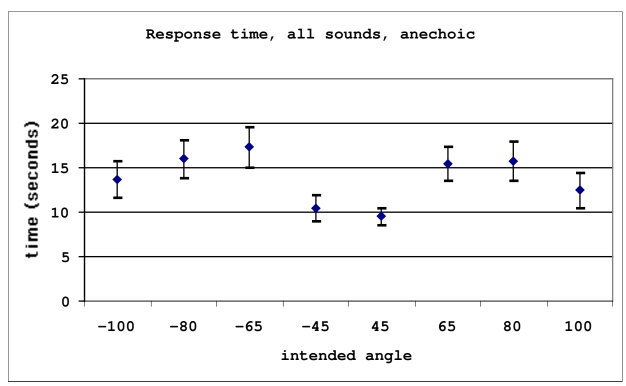

Fig. 24: Plot of mean response time (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the anechoic sound only condition.

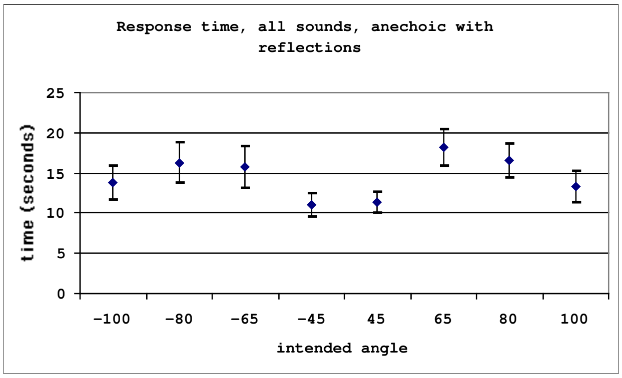

Fig. 25: Plot of mean response time (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the anechoic sound with reflections condition.

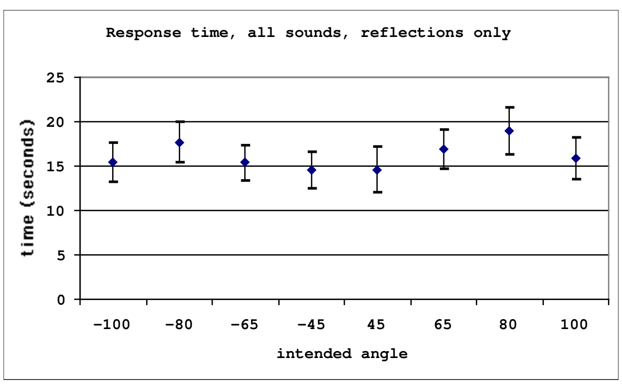

Fig. 26: Plot of mean response time (with 95% conf. int.) as a function of location for all three sound sources for all eight locations in the reflections only condition.