THE QUEEN'S AWARDS
FOR ENTERPRISE:
INNOVATION
2011

BBC

Research & Development

White Paper

WHP 261

October 2013

Localisation of Elevated Sources in Higher-Order Ambisonics

Paul Power, Chris Dunn, Bill Davies, Jos Hirst

BRITISH BROADCASTING CORPORATION

White Paper WHP 261

Localisation of Elevated Sources in Higher-Order Ambisonics

Paul Power, Chris Dunn, Bill Davies, and Jos Hirst

Abstract

Increasing the order of ambisonic reproduction allows 3D sound fields to be reconstructed with improved resolution and enhanced localisation, at the expense of higher channel counts, increased broadcast bandwidth, and increased storage requirements. Due to the resolution of human hearing it is possible that lower-order reproduction is sufficient for elevated sources in comparison to sources located in the horizontal plane, and so for 3D reproduction mixed-order ambisonic renderings may be the most efficient use of available transmission bandwidth and storage capacity. We report the results of subjective localisation tests for virtual sources placed in the vertical plane at various elevations and azimuths, for first, second and third-order ambisonic reproduction over a 16-loudspeaker system. The results provide insights into the requirements of higher-order ambisonics for broadcast and domestic reproduction.

The material in this paper was originally presented at the Institute of Acoustics *Reproduced Sound* conference in November 2012: P. Power, W. J. Davies, J. Hirst and C. Dunn, "Localisation of Elevated Virtual Sources in Higher-Order Ambisonic Soundfields," IOA conference *Reproduced Sound 28*, Brighton, UK (2012).

Additional key words: 3D sound, ambisonics, height, HOA, localisation, mixed order, surround sound

White Paper WHP 261

Localisation of Elevated Sources in Higher-Order Ambisonics

P Power University of Salford

C Dunn BBC Research and Development

W J Davies University of Salford

J Hirst University of Salford

1 Introduction

With current broadcast and consumer audio formats restricted to 2D horizontal-only spatial reproduction, improved listener experience may be possible with 3D technologies that reproduce height information. Higher-order ambisonics is one approach to recording, storing, transmitting and rendering 3D sound. Alternative non-ambisonic 3D rendering methods include VBAP (vector base amplitude panning) [1]. These and other rendering methods can be used with standardised 3D speaker layouts including NHK's 22.2 system [2], and Auro 3D [3].

Ambisonics was invented in 1969 by Michael Gerzon [4] and is essentially a two-stage system, where sound scenes can be recorded or synthetically encoded without knowledge of the reproduction setup. Reproduction resolution is dependent on the ambisonic order, and hence can be tailored to the accuracy required by using a higher- or lower-order coding.

This can be exploited for 3D reproduction since it is well known that human hearing resolution is less critical in the vertical plane than in the horizontal plane. There have been many studies to support this, including early work by Roffler and Butler [5] into the spatial localisation of tones at various frequencies. It was found that perceived sound source location depended on tone pitch rather than the real source location, furthermore to localise a sound displaced vertically requires frequency content above 7 kHz. Blauert [6] also characterised 'directional bands' where a sound's energy content at a specific frequency, rather than location, would dictate the perceived elevation of the sound. Due to these factors it may be possible to reduce the vertical ambisonic order compared to the horizontal order for a given localisation resolution. This leads to a mixed-order ambisonic system which may yield optimal reproduced quality for a given transmission bandwidth or storage capacity.

Several studies have examined the order requirements of 2D ambisonic systems, for example Frank *et al* [7] who show that in a reverberant room listeners could distinguish between the localisation performance of third- and fifth-order systems.

Research into the perceptual requirements of 3D surround systems has been limited, although a recent study by Capra *et al* [8] compared a triple stereo dipole system using a transaural approach to a sixteen-channel 3D ambisonic system using a first-order decoder optimised for the speaker array. It was found that there were no significant differences between the systems, and elevated sources were difficult to localise. Localisation confusions were also prevalent, in particular for sources placed below the participant which tended to be localised in the upper hemisphere.

Baume and Churnside [9] compared 2D and 3D first-order ambisonic systems in an attempt to understand the value of reproducing height information, and found that first-order with-height systems offer only small subjective improvements over 2D reproduction. Similarly no clear preference was found for 3D first-order ambisonics over a 5.1 surround system.

Morrell and Reiss [10] extended localisation comparisons to other methods including VBAP, time-delay panning, and third-order ambisonics for a 16-channel 3D speaker array. A variety of audio samples were used which were panned using the different methods. It was found that, using a variety of audio samples, the localisation performance of third-order ambisonics was better than the delay-based approach, and similarly to VBAP.

The speaker array used by Capra *et al* [8] was also used by Keiler and Batke [11] to compare two ambisonic decoding methods with VBAP, including a basic third-order ambisonic decoder using mode matching, and a proprietary 'robust' ambisonic method. Only two panned source positions were tested, and instead of directional estimates from participants a modified MUSHRA test regime using preference scores was adopted. It was found that the proprietary ambisonic panning method received the highest preference scores.

In this paper we study the localisation accuracy of 3D ambisonic systems for virtual sources placed in the vertical plane. Results are reported for systems up to third order, and allow some insight into order requirements for broadcast applications.

2 Ambisonic Reproduction

Ambisonics facilitates reproduction with scalable resolution, beginning at zero order which only accounts for the pressure component *W* of a soundfield at a single point in space. Adding 3D directional components (*X, Y, Z*) provides first-order directional information, allowing soundfield reproduction with limited spatial resolution. Accuracy can be improved by adding further components corresponding to higher orders. With each increase in ambisonic order the minimum number of loudspeakers required *A* also increases, where for ambisonic order *B,*

$$A = (B + 1)^2 \quad . \qquad (1)$$

The highest-order system investigated in this study is third order, requiring a minimum of 16 speakers.

Ambisonic reproduction assumes a regular speaker layout, which is easily realised for horizontally-panned sources since the speakers only have to be equally spaced. However for elevated sources this is more difficult to achieve as there are only five shapes - the platonic solids [12] - that provide regular speaker spacing.

A number of steps are required to reproduce a soundfield using ambisonics. Firstly the signal is encoded to the desired location, and is then decoded to a specific speaker array. For the encoding stage of our tests the Furse-Malham equation set was used with SN3D normalisation [13], where for source azimuth θ and elevation φ, the first-order components *W* to *Z*, second-order components *R* to *V*, and third-order components *K* to *Q* are given by

$$W = 1 \qquad (2)$$

$$X = \cos\theta \cos\varphi$$

$$Y = \sin\theta \cos\varphi$$

$$Z = \sin\varphi$$

$$R = (3 \sin^2\varphi - 1) / 2$$

$$S = \sqrt{3/2} \cos\theta \sin2\varphi$$

$$T = \sqrt{3/2} \sin\theta \sin2\varphi$$

$$U = \sqrt{3/2} \cos 2\theta \sin^2\theta$$

$$V = \sqrt{3/2} \sin 2\theta \cos^2\theta$$

$$K = \sin\varphi \, (5\sin^2\varphi - 3) / 2$$

$$L = \sqrt{3/8}\ \cos\theta\ \cos\varphi\ \left(5\sin^2\varphi - 1\right)$$

$$M = \sqrt{3/8}\ \sin\theta\ \cos\varphi\ \left(5\sin^2\varphi - 1\right)$$

$$N = \sqrt{15/2}\ \cos2\theta\ \sin\varphi\ \cos^2\varphi$$

$$O = \sqrt{15/2}\ \sin2\theta\ \sin\varphi\ \cos^2\varphi$$

$$P = \sqrt{5/8}\ \cos3\theta\ \cos^3\varphi$$

$$Q = \sqrt{5/8}\ \sin3\theta\ \cos^3\varphi\ \ .$$

A matrix of speaker directions *C* based on spherical coordinates is used to derive a pseudo-inverted speaker matrix *D.* When *D* is multiplied by the encoded signal direction $\vec{E}$ truncated to *B*th-order spherical harmonics, an approximation to the original soundfield is reconstructed:

$$D \cdot \vec{E_B} \simeq C^{-1} \cdot \vec{E} \qquad (3)$$

Bertet *et al* [14] explain that with higher-order ambisonics more selective use of the loudspeakers achieves higher sound source directivities. Decoder optimisation can be realised using coefficient weightings, for example 'max rE' (maximum energy) weighting, which aim to concentrate panned sources to the desired locations. Max rE decoding was used for the listening tests described in this paper.

3 Listening Test

A listening test was carried out to investigate the vertical localisation accuracy of ambisonic reproduction up to third order.

3.1 Physical set up

Three different resolutions of ambisonic decoding were used with the 16-speaker array configuration used by Capra *et al* [8] and Keiler and Batke [11]. Active Genelec 8030A speakers were arranged with eight placed on the horizontal plane at the listener's ear height, beginning at the centre-front 0 ° position and equally spaced in increments of 45 °. Another four speakers were arranged in a square above the listener, and four below, at ± 35 ° elevation offsets. The first speaker in each square layer was placed at an azimuth of 45 °, with further speakers placed at 90 ° azimuthal increments. The array radius was 1.35 m, with the listening position at the centre of the array (Fig. 1). An additional five speakers were placed to provide real sound sources.

A semi-anechoic chamber at the University of Salford was used for the tests, providing conformance to ISO 3744, ISO 3745, and BS 4196 standards. The working dimensions of this chamber are 4.2 x 3.3 x 3.0 m.

3.2 Sound sources

A critical aspect of listening tests is selection of appropriate audio samples since this can have a large bearing on the final results. Pink noise and speech have been used extensively in localisation tests by Keiler and Batke [11], Liebetrau *et al* [15], Barbour [16], and Naoe *et al* [17]. Carlile [18] explains that for accurate localisation spectral information across a wide range of frequencies is required, while Wightman and Kistler [19] suggest that when localising the direction of a broadband sound source the inter-aural time differences dominate. Liebetrau *et al* [15] note that humans are very sensitive to speech and such test signals should assist localisation. However Davis and Stephens [20] found that noise could be localised better than male speech in vertical localisation tests.

It would seem from previous work in the field that in terms of localisation there is no compelling case to exclusively use either pink noise or speech signals. Consequently our tests used both full-bandwidth pink noise bursts, and a female speech sample taken from the music test CD *Music for Archimedes* [21]. Each of the 44.1 kHz 16-bit samples was edited to two seconds in length.

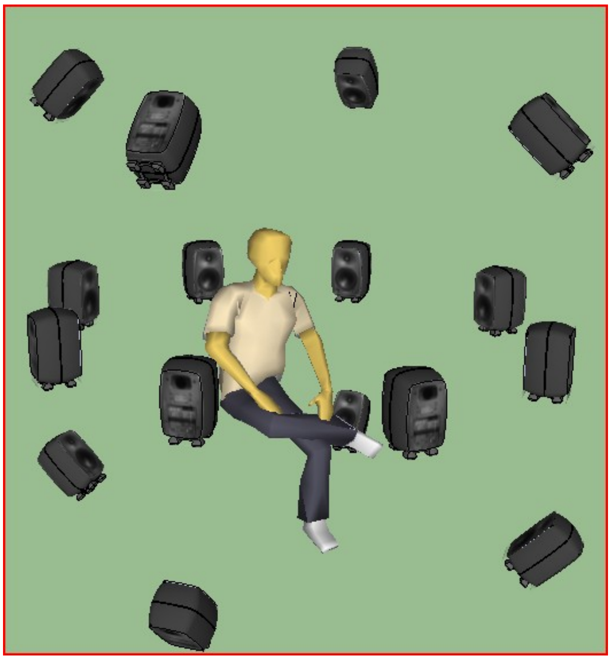

Fig 1. Ambisonic array with 16 speakers used to pan virtual sources in listening tests. Five further speakers (not shown) were used to provide real source locations.

3.3 Experimental design

The two sound samples were processed using first, second and third-order ambisonics. In addition to five real sound sources, eight virtual sources were used for each order. Hence each participant test included a total of fifty eight sound sources, split evenly between noise and speech samples (Table 1).

Source location		**Virtual source**	**Real source**
Azimuth º	**Elevation φ**		
20	35	✔	✔
20	-35	✔	
70	20	✔	
70	35	✔	✔
90	35	✔	✔
300	0	✔	✔
330	-35	✔	
330	35	✔	✔

Table 1. Real and virtual source locations. Localisation accuracy was assessed at each source location using pink noise and speech samples.

The sound samples were replayed using computer software interfaced to a multichannel digital-to-analogue converter via MADI (multichannel audio digital interface). The audio samples were arranged in a random order for each participant. Assigning cue points in the playback software allowed the experimenter to control instant switching between samples.

An acoustically-transparent curtain was hung within the semi-anechoic chamber in order to obscure the frontal speakers, however the rear loudspeakers were visible to participants on entering the chamber. No speakers were visible while seated at the listening position.

Each of the active monitors in the 16-speaker array were adjusted in level using pink noise and a sound level meter to produce 70 dBA at the listening position. Further level alignment was made by normalising first, second and third-order system outputs measured at the listening position, to the front centre channel only of the horizontal array, using pink noise excitation. The level adjustments required for equal output were -9.8 dB for first and second order, and -5.2 dB for third order. The levels for each of the real sources replayed from individual speakers were also adjusted using pink noise for 70 dBA at the listening position.

3.4 Recording participant responses

Collecting participant responses in 3D localisation experiments is not a trivial process, and several methods have been used by others working in the field. Pernaux *et al* [22] evaluated different methods of reporting 3D sound source positioning, and found that finger pointing with a 3D visual interface provided the least error. Wightman and Kistler [23] found that verbally reporting source azimuth and elevation in spherical coordinates gave the same results as using a clock face method. Other methods were also explored by Evans [24]. With the resources available for our experiments it was decided to verbally report the perceived sound source positions in spherical coordinates, using an intercom system that allowed communication between operator and participants. To assist participants in their judgements, a vertical elevation scale was printed on the curtain hung in front of listeners. The scale had tick marks at 5 º increments, with 0 º being directly straight ahead at the same height as the centre-front speaker tweeter. The elevation scale spanned a range of ± 55 º, and was calibrated using an inclinometer. The limited curtain radius did not allow a similar azimuth scale to be used, instead a sheet of paper with azimuth markings was used.

Participants were told that the two-second audio samples would be played three times, however they could request samples to be repeated if required. After the participants had been provided with instructions and allowed to ask questions the test was started.

Eleven participants took part, and were either staff or postgraduate researchers of the Acoustic Research Centre at the University of Salford. Since all participants were either studying or researching audio technology they were classed as selected assessors [25].

4 Results

For each participant, azimuth and elevation responses were recorded for each source location. Elevation error was calculated as the magnitude of the difference between a participant's response and the intended panned or real location. Azimuth errors were also calculated but are not reported in this paper.

In order to evaluate elevation errors for virtual sources, it was necessary to first assess performance with respect to real sources. Strybel and Fujimoto [26] explain that the minimum audible vertical angle resolution is approximately 4 to 5 º, however vertical resolution is reduced at higher elevations. Inspecting elevation errors for the real source positions shown in Fig. 2 it can be seen that the average elevation error is slightly greater than 5 º. Using a different reporting method from the approach used in our tests may achieve greater accuracy.

For the virtual source localisation results shown in Figs. 3 - 5, significant variation in elevation error is observed for the three ambisonic orders tested. The large confidence intervals recorded for all orders with sources placed below the participant, in particular at (20º, -35φ), were caused by up / down and front / back confusions. However the number of confusions was significantly lower for the third-order system (Fig. 6). Where participants had up / down confusions a number of front / back confusions were also observed (Table 2).

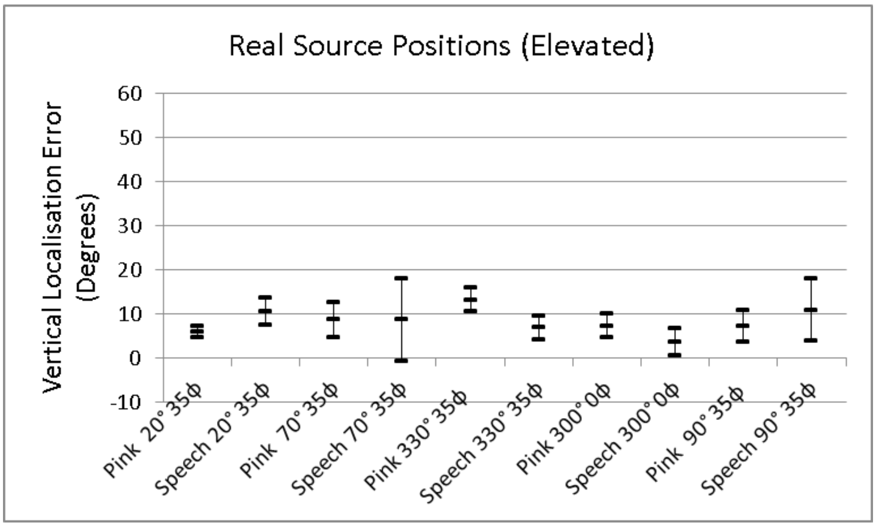

Fig. 2. Elevation errors for real sources, including mean and 95% confidence intervals.

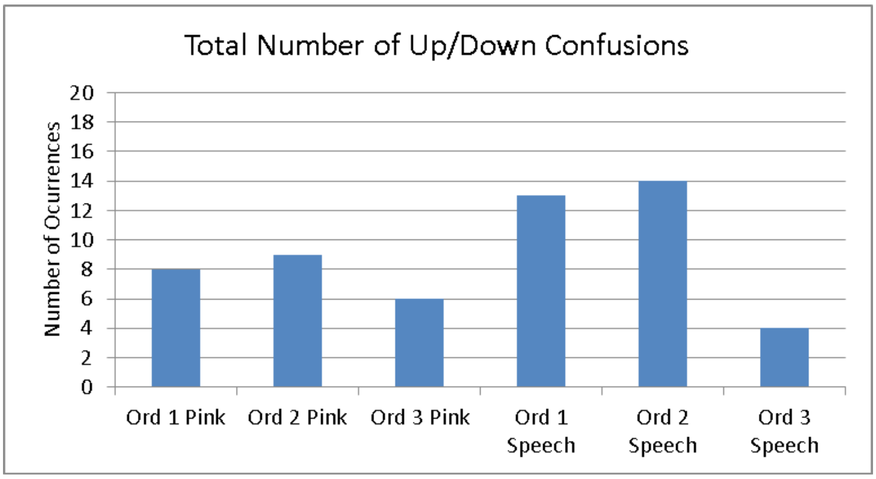

Fig. 6. Dependency of up / down confusions on ambisonic order and test material.

Test signal	**Ambisonic order**	**Virtual source position**					**Total confusion count**
		20º 35φ	**20º -35φ**	**300º 0φ**	**330º 35φ**	**330º -35φ**	
Female speech	1	3	3	0	4	2	12
	2	1	4	1	2	2	10
	3	0	3	0	0	1	4
Pink noise	1	1	4	0	1	4	10
	2	2	2	5	2	4	15
	3	0	3	0	0	0	3

Table 2. Dependency of front / back confusion frequency on ambisonic order, virtual source position, and test material.

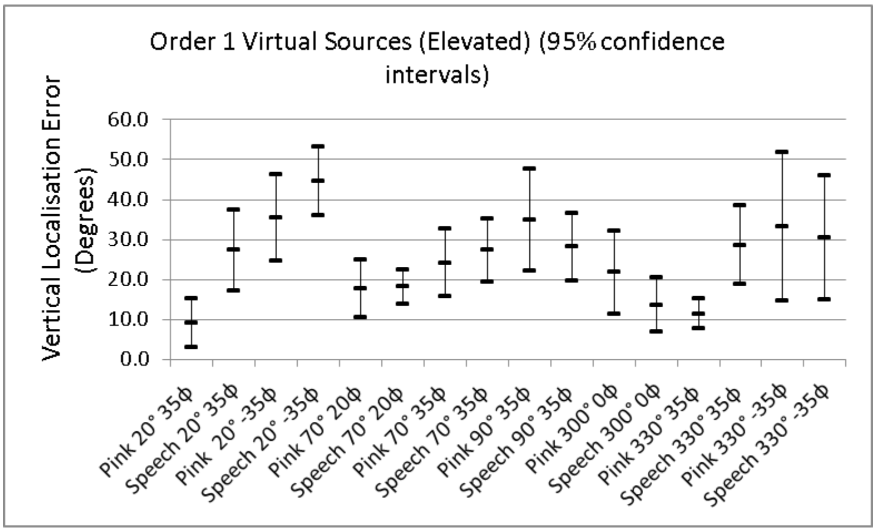

Fig 3. Elevation errors of virtual sources reproduced by first-order ambisonic system, including mean and 95% confidence intervals.

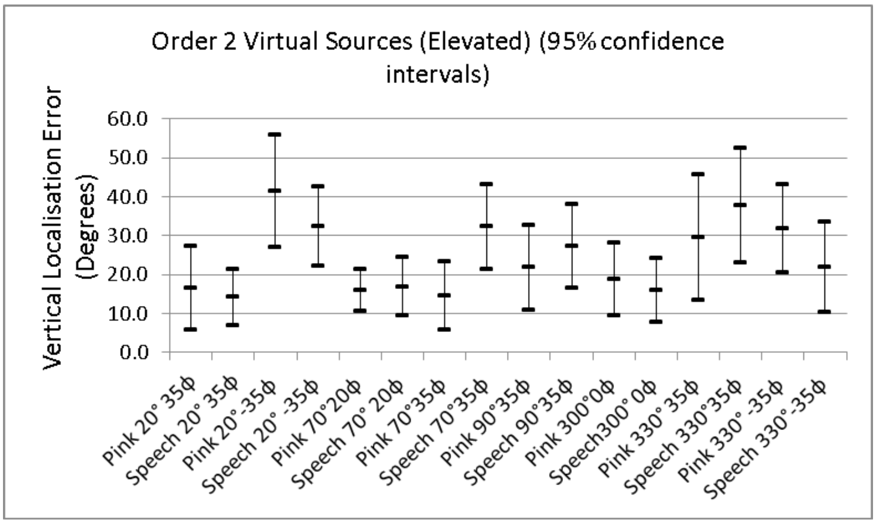

Fig 4. Elevation errors of virtual sources reproduced by second-order ambisonic system, including mean and 95% confidence intervals.

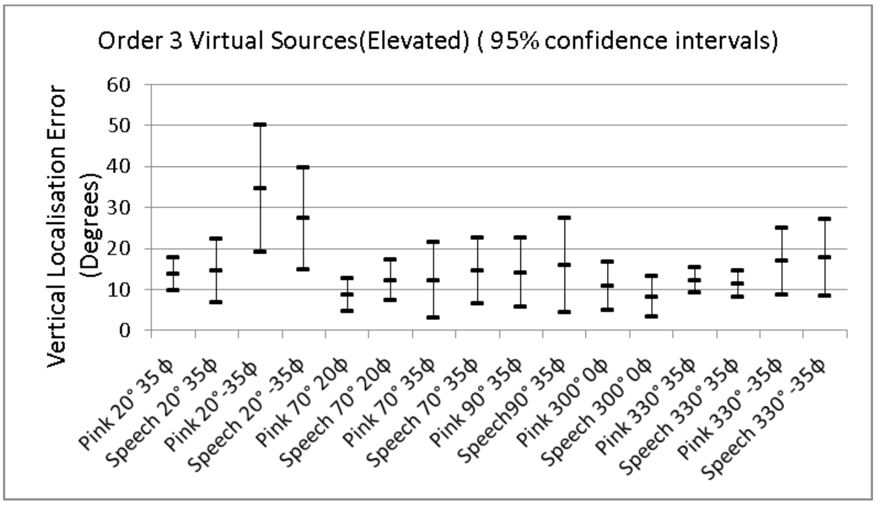

Fig 5. Elevation errors of virtual sources reproduced by third-order ambisonic system, including mean and 95% confidence intervals.

To determine which ambisonic order performed with the least elevation error, the mean error for all elevated source positions was examined. Fig. 7 indicates that for the speech item, the third-order system has a lower average error than the first- and second-order systems. A similar trend is also observed for the pink noise item, however here the differences are not statistically significant.

In order to find if the performance of the third-order rendering was significantly different to the lower-order systems for speech or pink noise, a one way ANOVA was carried out at the 0.5% significance level [27]. For the pink noise item no significant difference could be found between orders for elevated virtual sources [$F(2,21) = 2.032$, $p > 0.05$]. However for the speech sample there were significant differences between orders [$F(2,21) = 5.17$, $p < 0.05$].

To determine any specific differences between ambisonic orders a Bonferroni post hoc test [27] was carried out. The results shown in Table 3 indicate a significant difference between first and third orders in terms of overall mean localisation error, with third order having a significantly smaller error. However there was no significant difference between second- and third-order systems.

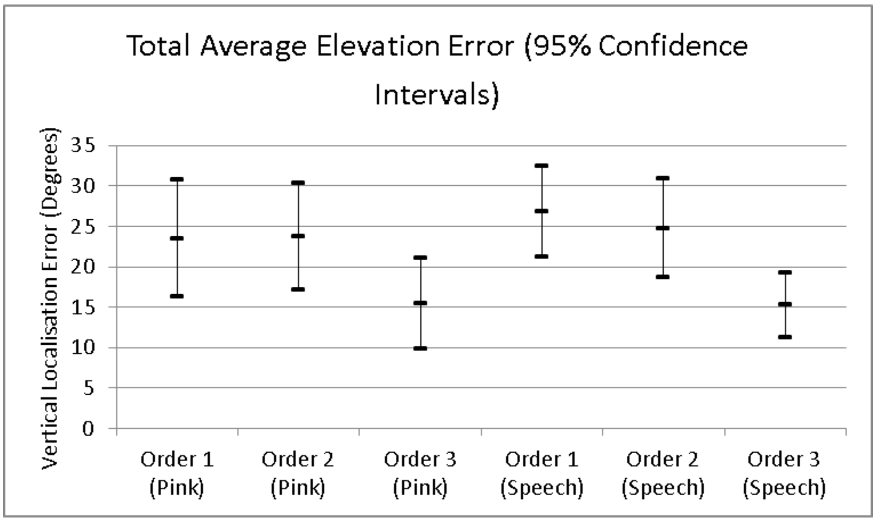

Fig 7. Dependency of average elevation error on ambisonic order and test material for elevated source positions, including mean and 95% confidence intervals.

(I) order	**(J) order**	**Mean difference (I - J)**	**Std. error**	**Sig.**	**95% confidence interval**	
					Lower bound	**Upper bond**
order 1	order 2	2.04500	3.84654	1.000	-7.9612	12.0512
	order 3	11.59000*	3.84654	.020	1.5838	21.5962
order 2	order 1	-2.04500	3.84654	1.000	-12.0512	7.9612
	order 3	9.54500	3.84654	.065	-.4612	19.5512
order 3	order 1	-11.59000*	3.84654	.020	-21.5962	-1.5838
	order 2	-9.54500	3.84654	.065	-19.5512	0.4612

* The mean difference is significant at the 0.05 level.

Table 3. Bonferroni multiple comparisons of mean elevation error between different ambisonic orders.

5 Discussion

The listening test results clearly show the inconsistency of first- and second-order ambisonic systems in localising virtual sources, compared to the third-order system which localised sources with the least error for both elevation and azimuth judgements. The poor localisation achieved with first-order systems may partly explain the limited subjective improvements observed in previous comparisons of first-order with-height ambisonics with 2D ambisonics and 5.1 surround systems [9].

No statistically significant differences in localisation accuracy could be established between different ambisonic orders for elevated sources using the pink noise sample (Fig. 7). In contrast significant differences were found between orders for the speech item which has a more limited frequency range, and also provides transient cues. Liebetrau *et al* [15] note that human hearing is very sensitive to speech, which could make participant localisation of panned sources more critical. Larger azimuth errors were also found for first and second order, behaviour that could be due to angular dispersion at higher frequencies [28].

For virtual source positions constrained to the horizontal participants tended to perceive sources displaced from the horizontal, this effect was more pronounced for first-order renderings but was also evident for real sources. Larger offsets were more evident for the pink noise source than speech, possibly due to more high frequency content. Overestimation of elevated sound sources was also found by Liebetrau *et al* [15].

In our tests elevated sources placed in front of participants caused a number of front / back confusions with the first- and second-order systems, while for third order front / back confusions only occurred when sources were placed below participants. The close proximity of the lower speaker layer to the listener may contribute to the up / down confusions experienced. Up / down confusions were also found by Capra *et al* [8] for a first-order system, while in our tests fewer up / down confusions were observed with the third-order material. The relatively high frequency of confusions experienced with first and second order may be due to the large number of speakers that simultaneously contribute to these system outputs, in contrast to the third-order case where fewer speakers are simultaneously used to pan virtual sources [14].

There were a number of non-ideal aspects to our test design. Firstly the method used to report participant responses, based on participants translating source location into spherical coordinates, could have inflated localisation errors for both real and virtual sources. A second significant issue is that the speaker array used for our tests was irregular. With eight speakers placed on the horizontal and four above and below, sound sources on the horizontal will tend to advantage the third-order system as eight speakers is greater than the number required for optimal first- and second-order reproduction [29]. Related arguments suggest that for elevated sources the speaker array used for our tests will tend to disadvantage second and third orders.

6 Conclusions

Ambisonic reproduction systems of first, second and third orders using a 3D speaker array were evaluated in terms of vertical localisation accuracy in a semi-anechoic environment with two audio samples.

It was found that vertical localisation accuracy was inconsistent for the first- and second-order systems, with third order providing more consistent results. Significant differences in localisation accuracy were found between first- and third-order systems for a speech sample, although with pink noise no statistically-significant difference was found between orders. This suggests the optimal ambisonic order for reproducing elevated sound sources is dependent on the desired accuracy of localisation and the frequency content of the sound source.

In general the results indicate that third-order ambisonics is more effective at accurately localising elevated panned sources than lower-order systems, although even third-order localisation accuracy remains below the localisation resolution of human hearing indicated by the results for real sources. For elevated sources, greater than third-order ambisonics may be required to match the localisation resolution of human hearing.

It should be noted that our experimental results were obtained under semi-anechoic conditions, and in real-world reverberant environments there may be a reduced ambisonic order requirement for the vertical plane.

7 References

[1] V. Pulkki, “Virtual Sound Source Positioning using Vector Base Amplitude Panning,” *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456 - 466 (1997 June).

[2] K. Hamasaki, K. Hiyama, and R. Okumura, "The 22.2 Multichannel Sound System and its Application," presented at the 118th Convention of the Audio Engineering Society, paper 6406 (2005 May).

[3] B. Van Daele and W. Van Baelen, "Productions in Auro 11.1", Barco White Paper (2012).

[4] M. F. Davis, "History of Spatial Coding," *J. Audio Eng. Soc.*, vol. 51, no. 6, pp. 554 - 569 (2003 June).

[5] S. K. Roffler and R.A. Butler, "Factors that Influence the Localization of Sound in the Vertical Plane," *J. Acoust. Soc. Am.*, vol. 43, no. 6, pp. 1255 - 1259 (1968).

[6] J. Blauert, *Spatial Hearing* (MIT Press, 1996).

[7] M. Frank, F. Zotter, and A. Sontacchi, "Localization Experiments using Different 2D Ambisonics Decoders," 25th Tonemeistertagung - VDT Int. Convention (2008 November).

[8] A. Capra et al., "Listening Tests of the Localization Performance of Stereo Dipole and ambisonic Systems", presented at the 123rd Convention of the Audio Engineering Society, paper 7187 (2007 Oct.).

[9] C. Baume and A. Churnside, "Scaling New Heights in Broadcasting using Ambisonics," Proc. 2nd Int. Symposium on Ambisonics and Spherical Acoustics, Paris, France (2010 May).

[10] M. J. Morrell and J. D. Reiss, "A Comparative Approach to Sound Localisation within a 3D Sound Field," presented at the 126th Convention of the Audio Engineering Society, paper 7663 (2009 May).

[11] F. Keiler and J. M. Batke, "Evaluation of Virtual Source Localization Using 3-D Loudspeaker Setups", presented at the 128th Convention of the Audio Engineering Society, paper 8060 (2010 May).

[12] F. Hollerweger, "Periphonic Sound Spatialization in Multi-User Virtual Environments", Masters thesis, Institute of Electronic Music and Acoustics, Graz, Austria (2006).

[13] D. Malham, "3-D Acoustic Space and its Simulation using Ambisonics," *Computer Music Journal*, vol. 19, no. 4, pp. 58 - 70 (1995).

[14] S. Bertet, J. Daniel, and S. Moreau, "3D Sound Field Recording with Higher-Order Ambisonics - Objective Measurements and Validation of Spherical Microphone," presented at the 120th Convention of the Audio Engineering Society, paper 6857 (2006 May).

[15] J. Liebetrau et al., "Localization in Spatial Audio - from Wave Field Synthesis to 22.2", presented at the 123rd Convention of the Audio Engineering Society, paper 7164 (2007 Oct.).

[16] J. Barbour, "Elevation Perception: Phantom Images in the Vertical Hemisphere," Proc. Australasian Computer Music Association Conference (2004 July).

[17] M. Naoe et al., "Performance Evaluation of 3D Sound Field Reproduction System Using a Few Loudspeakers and Wave Field Synthesis," IEEE Int. Symp. on Universal Communication, ISUC '08, Osaka, Japan (2008 Dec.).

[18] S. Carlile, *Virtual Auditory Space: Generation and Applications* (R G Landes, Austin, Texas 1996).

[19] F. L. Wightman and D. J. Kistler, "The Dominant Role of Low-Frequency Interaural Time Differences in Sound Localization," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1648 - 1661 (1992 Mar.).

[20] R. Davis and S. Stephens, "The Effect of Intensity on the Localization of Different Acoustical Stimuli in the Vertical Plane," *Journal of Sound and Vibration*, vol. 35, no. 2, pp. 223 - 229 (1974 July).

[21] Bang & Olufsen, *Music for Archimedes* (1992).

[22] J. M. Pernaux, M. Emerit, and R. Nicol, "Perceptual Evaluation of Binaural Sound Synthesis: the Problem of Reporting Localization Judgements", presented at the 114th Convention of the Audio Engineering Society, paper 5789 (2003 Mar.).

[23] F. L. Wightman and D.J. Kistler, "Headphone Simulation of Free-Field Listening: II. Psychophysical Validation," *J. Acoust. Soc. Am*., vol. 85, no. 2, pp. 868 - 878 (1989 Feb.).

[24] M. J. Evans, "Obtaining Accurate Responses in Directional Listening Tests," presented at the 104th Convention of the Audio Engineering Society, paper 4730 (1998 May).

[25] S. Bech and N. Zacharov, *Perceptual Audio Evaluation* (John Wiley & Sons, London, 2006).

[26] T. Z. Strybel and K. Fujimoto, "Minimum Audible Angles in the Horizontal and Vertical Planes: Effects of Stimulus Onset Asynchrony and Burst Duration," *J. Acoust. Soc. Am.*, vol. 108, no. 6, pp. 3092-3095 (2000 Dec.).

[27] A. P. Field, *Discovering Statistics using SPSS* (Sage Publications Limited, London, 2009).

[28] J. Daniel, "Evolving Views on Higher-Order Ambisonics: from Technological Views to Pragmatic Concerns," Proc. Ambisonics Symposium, Graz, Austria (2009 June).

[29] A. Solvang, "Spectral Impairment of Two-Dimensional Higher Order Ambisonics," *J. Audio Eng. Soc.*, vol. 56, no. 4, pp. 267 - 279 (2008 April).