THE QUEEN'S AWARDS FOR ENTERPRISE: INNOVATION 2011

Research & Development

White Paper

WHP 221

March 2012

Upping the Auntie

A broadcaster's take on ambisonics

C. Baume ***and*** **A. Churnside**

BRITISH BROADCASTING CORPORATION

BBC Research
White Paper WHP 221

Upping the Auntie

Chris Baume and Anthony Churnside

Abstract

This paper considers Ambisonics from a broadcaster's point of view: to identify barriers preventing its adoption within the broadcast industry and explore the potential advantages were it to be adopted. This paper considers Ambisonics as a potential production and broadcast technology and attempts to assess the impact that the adoption of Ambisonics might have on both production workflows and the audience experience. This is done using two case studies: a large-scale music production of "The Last Night of the Proms" and a smaller scale radio drama production of "The Wonderful Wizard of Oz". These examples are then used for two subjective listening tests: the first to assess the benefit of representing height allowed by Ambisonics and the second to compare the audience's enjoyment of first order Ambisonics to stereo and 5.0 mixes.

Additional key words: spatial, Soundfield, B-format, perceptual

White Papers are distributed freely on request.

Authorisation of the General Manager is required for publication.

This paper was originally published at the 128th Audio Engineering Society Convention in London, May 22–25 2010.

1 Introduction

“Delivering to the public the benefit of emerging communications technologies and services” [1] is one of the six public purposes of the British Broadcasting Corporation. This statement explains why BBC Research and Development explores new technologies and assesses their potential benefits to the public. The BBC broadcasts over 400 hours of audio content every day, which is why investigating the cost saving and quality enhancing benefits of developments in audio technology is an important part of BBC R&D’s work. Whilst the development of Ambisonics began in the 1970s, the availability of fast and low-cost digital processing, combined with freely-available digital production tools, means that it is now much easier for broadcasters to produce Ambisonic content.

2 Theoretical Benefits

Ambisonics offers a number of theoretical benefits to the BBC and its audience. A challenge faced by the BBC is how to deal with the various channel formats in which audio content is produced. The compatibility of mono, stereo and 5.1, which of them to broadcast and which to archive are all important questions. Ambisonics could provide a solution to these problems. If a B-Format [2] (or high order Ambisonics) representation of the sound field were used in place of the mono, stereo and 5.1, this might provide compatibility with current formats and could even help future proof the archive.

Another possible advantage could be the speaker-agnostic nature of a sound-field representation. It is problematic for the audience who listen to BBC content in 5.1 to set-up their listening environment to ITU [3] standards. It is likely that most achieve only an approximation to the recommendation, at best. Ambisonic content can be decoded for a wide variety of speaker numbers and layouts, which would allow individual audience members to configure their listening environment in a way that best suits their requirements, without compromising the quality of reproduction. This would also help avoiding such problems as a centre speaker position needing to share the same space as a screen.

Finally, an obvious advantage of Ambisonics is its ability to convey height information. As will be discussed later, this may offer an improved listening experience. However it can also act as a useful marketing tool, as it can be considered to be a ‘3D’ audio format.

3 Case Studies

Two case studies have been undertaken to assess the use of Ambisonics against current formats used in ‘typical’ BBC productions. It was thought important to learn how the use of Ambisonic formats might impact on current workflows. Two productions were selected for their contrasting workflows; “The Last Night of the Proms” [4], a high profile, live outside broadcast (OB); and “The Wonderful Wizard of Oz” [5], a smaller scale pre-recorded Radio 4 drama.

3.1 Live Music: The Last Night of the Proms

The Last Night of the Proms is a flagship BBC production which is broadcast live every year. It is the last evening of a season of eight weeks of classical music concerts, the majority of them held at the Royal Albert Hall in London. This event was chosen as a useful case study for a number of reasons. Firstly, it is a live broadcast of a large scale and high complexity. Secondly, it takes place in a large spacious environment and the aim of the producer is to accurately represent the sound at the venue.

3.1.1 Capturing the Proms

The microphone placement in the hall for the Proms remains fairly constant with only minor changes made from year to year. It uses a mixture of spot microphones close to the different sections of the orchestra and the soloists, curtains of cardioid microphones hung across the stage, traditional Decca-tree [6] microphone arrangements and rear facing arena microphones (see figure 1). The total number of microphones used tends to exceed 120. Over 30 were hung on slings between the left and right upper circle, about 20 were close microphones split with the Royal Albert Hall sound reinforcement (PA) system and over 40 were close microphones not sent to the PA. This may seem like a large number of microphones, but it is important to highlight that the venue is rigged once for the full eight week season and the set-up must cope with a very eclectic mix of classical performances. Not all the microphones are used for all the performances.

The basis of the 5-channel surround mix was made from the combination of a Decca tree for the front left, centre and front right. An arrangement known as the "Hamasaki square" [7] provided more reverberant sound for the rear left and right channels of the 5.1 mix.

The basis for the B-Format recording of The Last Night of the Proms was a Soundfield DSF-2 Digital Broadcast Microphone System [8]. The microphone was located approximately 3 meters behind and 5 meters above the conductor's position. The microphone was tilting down approximately 45° to point towards the centre of the orchestra. The B-format digital outputs from the DSF-2 were captured using two synchronised soundcards and a laptop computer. The rest of the microphone signals were also captured using a Pyramix system.

3.1.2 Producing the Proms

The Last Night of the Proms was a live production where the stereo and 5.1 mix were made in an OB truck parked next to the Royal Albert Hall. The microphones detailed in section 3.1.1 were balanced in a 5.1 listening environment in the OB truck, which was equipped with a large digital desk (see figure 2). The engineer produced both the stereo and 5.1 mix at the same time. The 5.1 mix used the microphones from the Hamasaki square which were panned appropriately to the rear, whilst Decca trees were panned to the front left, centre and front right channels as the basis for the mix. Point sources were then appropriately panned to reinforce the image provided by the Decca trees and Hamasaki square. The Ambisonic mix was produced at a later date using a powerful laptop computer. The B-format from the DSF-2 was used as the basis of the Ambisonics mix and the close microphones from the multitrack recording were added to reduce the relative reverberant qualities of the DSF-2 output. To mix in these mono sources Steinberg's Nuendo software was

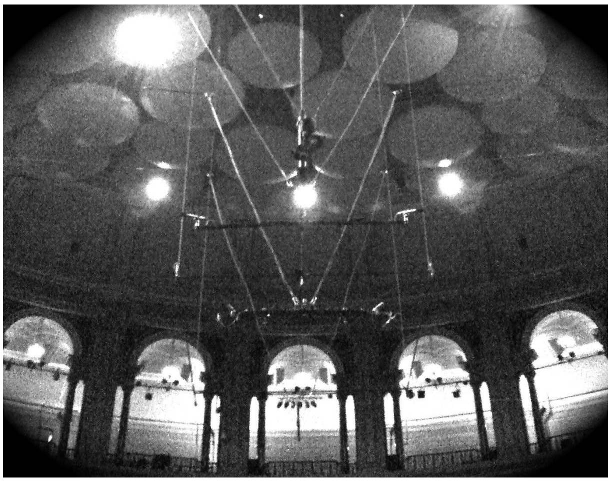

Figure 1: View from the stage of some of the microphones used to capture The Last Night of the Proms. Note a number of Decca tree microphone arrays are hung from different slings

Figure 2: The OB truck where the stereo and 5.1 mix of The Last Night of the Proms took place. The signals to the rear speakers (top right foreground of the photograph) are appropriately delayed to simulate a ITU-R BS.1116-1 5.1 set-up

used. The close mic'ed sources were virtually placed, relative to the position of the Soundfield microphone and the levels were mixed to resemble the balance of the 5.1 mix that was made live on the night. The process of recreating the sound field was somewhat similar to 5.1, but with the added height dimension. A 3D model of the listening environment used for playback of the Ambisonics mix is shown in figure 3.

3.2 Radio Drama: The Wonderful Wizard of Oz

Central to the production of good radio drama is the ability to create convincing artificial environments in which to place the narrative. This production of The Wonderful Wizard of Oz was selected as a case study for its differences to The Last Night of the Proms. This radio drama was set in a fictional world so it required the creation rather than recreation of spaces. Many of the performances were recorded in an acoustically dead space, unlike the Royal Albert Hall. Another major difference is that it was not broadcast live, it was recorded over the course of 3 days in a radio drama studio at BBC New Broadcasting House in Manchester.

3.2.1 Capturing The Wonderful Wizard of Oz

The production was recorded in non-chronological order, where actors often performed the same scene more than once, allowing the producer to select the best performance. The capture of the actors' performances for the stereo and 5.1 mixes used a stereo microphone (cardioid cross-pair). The actors generally performed in one of two acoustically distinct spaces: an acoustically 'dead' room to simulate being outside (or to allow reverberation to be added later) and an acoustically 'live' room which had acoustic qualities that conveyed that the performance was taking place inside a room. The output of the stereo microphone was captured to a Macintosh Computer running "Pro Tools". Diegetic sound effects were added, using a mixture of effects captured by a foley artist [9] using a stereo microphone, and stereo files from a sound effects archive. The music was written by a freelance composer and supplied as stereo audio files.

For the Ambisonic recording, the actor's performances were captured using an ST250 Soundfield microphone that was placed alongside the stereo microphone. The producer of the programme was aware of the additional creative possibilities presented by the use of Ambisonics in terms of the ability to represent three dimensional space. As a result extra direction was given to the performers to use the space around the Soundfield Microphone. While this provided opportunities for the producer to convey their picture of the drama, the Soundfield microphone recording suffered from structure-borne vibration caused by the movement of the actors stepping around the microphone stand. The B-format output of the ST250 was captured by a soundcard connected to a laptop computer.

3.2.2 Producing The Wonderful Wizard of Oz

The Ambisonics mix of The Wonderful Wizard of Oz was carried out in the same way as the Proms Ambisonics mix, using a powerful laptop computer running Steinberg's "Nuendo" software. The mix took place in the listening room shown in figure 3. The sound effects used were a combination of foley effects and archived sound effects, which were treated as mono sources and placed using dynamic Ambisonic panning.

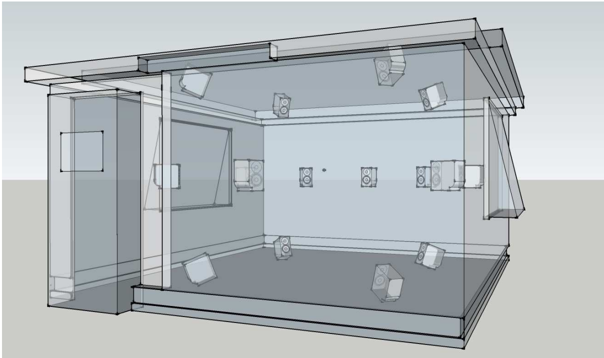

Figure 3: The listening room with the combined Ambisonics, 5.0 and stereo loudspeaker arrangement

The majority of the dialogue used in the Ambisonics mix was captured by the ST250, but on occasions when the recordings were affected by structure-borne noise as described above, the mono recordings of the dialogue were virtually placed and encoded into a ‘dry’ B-format.

To provide a realistic room response, impulse responses in a number of acoustic environments were captured in B-format, as shown in figure 4. These responses were convolved with the ’dry’ B-format containing the virtually placed mono sources. This B-format convolution technique was also used to mix in ’dry’ sound effects without them sounding acoustically out of place with the rest of the audio.

3.3 Conclusions from the Two Case Studies

Ambsonics appears to lend itself well to the reproduction of a live event such as The Last Night of the Proms, although there seems to be a limited number of B-Format production tools available at present. Simultaneously playing back and processing a large number of B-format proved taxing for the hardware, and often tested the limits of the computer’s CPU. Mixing mono sources with a B-format recording leads to the computer having to deal with a large number of tracks. For example, to mix the Proms in first order Ambisonics using 120 mono sources meant the computer had to mix a total of 480 tracks down to 4.

With The Wonderful Wizard of Oz the process was similar and therefore comparable problems were encountered. For this recording, much more post-production was required. Whilst techniques such as impulse response convolution did provide excellent results, the process was often tedious and the software tools available were not found to be particularly suited to the creative workflow.

4 Subjective Tests

In an attempt to assess the audience experience of B-Format material, a series of listening tests were carried out. A variety of listeners were asked to rate their enjoyment of listening experiences. Whilst quantitative data was collected, it was also considered important to collect detailed qualitative data. Both data sets are presented here.

Figure 4: Capturing the B-Format impulse response of a reverberant space

4.1 The Value of Height

Although Ambisonics is much more than a method of recording and replaying 'with-height' audio, it is often cited as a '3D' audio format. There is no doubt that the ability to include height is one of the best known examples of Ambisonics, and it gets people interested and excited about the technology. However, the effect that height information has on the listening experience is not currently well-understood. To assess the value of 3D audio, the effects of the representation of height need to be investigated.

In order to do this, a subjective test was designed in which several B-format clips were replayed using a variety of speaker layouts, some of which included speakers above and below the listener. Although Ambisonics was used to record and replay the test items, the intent was not to test the performance of Ambisonics itself.

4.1.1 Setup

A listening room in BBC R&D's previous base at Kingswood Warren in Surrey was used to conduct the listening test. Twelve PMC DB1-SA active monitors were used in the layout shown in Table 1. Six of the speakers were arranged in a hexagon layout in the horizontal plane, with three arranged in a triangle layout above the listener, and three in a triangle below (rotated 180°). This set up is shown in figure 5.

The twelve speakers were used in five different configurations, but for the purposes of this paper the results from three will be presented:

- Hex - Hexagon of speakers in the horizontal plane
- HexTri - All of the speakers, consisting of the hexagon in the horizontal plane, and the triangles above and below the listener.
- HexTriNoZ - All of the speakers, consisting of the hexagon in the horizontal plane, and the triangles above and below the listener, but with the Z channel (the height information) omitted.

Decoding matricies were generated for each configuration, which ensured that the overall sound

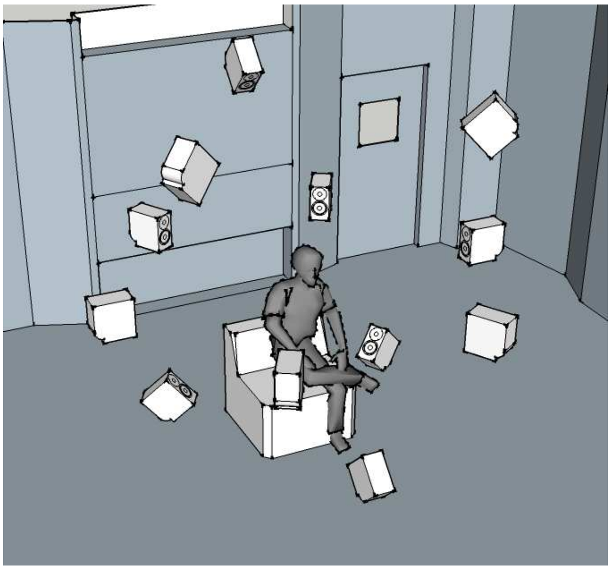

Figure 5: 3D model of the speakers used for the listening test

Speaker	X	Y	Z	Azi	Ele
1	Front	-	Down	0°	-45°
2	Back	Left	Down	120°	-45°
3	Back	Right	Down	240°	-45°
4	Front	Left	-	30°	0°
5	-	Left	-	90°	0°
6	Back	Left	-	150°	0°
7	Back	Right	-	210°	0°
8	-	Right	-	270°	0°
9	Front	Right	-	330°	0°
10	Front	Left	Up	60°	+45°
11	Back	-	Up	180°	+45°
12	Front	Right	Up	300°	+45°

Table 1: Speaker positions for the first listening test

level would be fairly consistent across all of them. Maximum velocity decoding was used for frequencies below 400Hz, and maximum energy decoding was used for frequencies above.

4.1.2 Method

The test was conducted using the MUSHRA [10] test method. Four different speaker layouts (or configurations) were considered, plus a hidden reference and an ‘anchor’ (which was the HexTriNoZ configuration). Each of these six configurations was given a score of how it compared to a given reference configuration. Five audio test items were used for the test, each one being played in a 30-second loop while the participant rated each configuration.

To speed up the test process, a user interface was designed to let the test participants dynamically control the speaker configuration in use. The GUI was modelled on examples in the MUSHRA recommendation, and was implemented using Java and Swing. The interface could be controlled using a keyboard and/or a mouse, and allowed users to select the active speaker configuration, as well as give a score for each one. The user’s actions and final results were saved locally in spreadsheets for later analysis. The software worked by sending MIDI messages to another PC running Steinberg Nuendo, which was used to play the audio.

The verbal scoring system used for the test was 'Much better', 'Slightly better', 'About the same', 'Slightly worse' and 'Much worse'. To line up with the verbal score, a numeric scale of $+20$ to -20 was used.

4.1.3 Items

Five separate audio test items were used in the listening test. Selecting suitable items was one of the most difficult parts of the experiment, as it can greatly influence the results. A variety of music and atmospheric items were chosen, with only two of the items containing explicit audio sources above the listener, the rest relying on reverberation for height content. Each of them is described below:

1. Classical music and applause: The BBC Concert Orchestra at the Proms 2009
   The raw B-format output of the Soundfield microphone system was used, containing a clip of classical music followed by applause.

2. Ambient soundscape: Thunder
   This item was recorded using a Soundfield ST-250 microphone at R&D's previous base at Kingswood Warren during a thunder storm. The microphone is held under an umbrella, so there is significant height content from the rain hitting the umbrella directly above.

3. Radio drama: The Wonderful Wizard of Oz.
   A very dramatic section where Dorothy's house is swept up into a tornado. It features some dialogue and a lot of sound effects.

4. Classical music
   This clip was taken from Ambisonia.com, and was made by Aaron J Heller. The recording is of an orchestra playing Beethoven's Symphony No. 4 in B-flat major, and was done using a Calrec Soundfield MkIV No. 99. Although there is noticeable reverberation, the orchestra sounds very close and is much 'drier' than the Proms recording. There is also more bass content, and unlike the Proms clip, it contains no applause.

5. Proms atmosphere
   Taken from the same recording as the 'Classical music and applause' clip, this does not contain music, but rather only audience noise. This includes clapping, laughing and horns from around the venue. The reverberation of the space is very noticeable in this clip.

4.1.4 Results

18 people took part in the listening test, 6 of whom had heard periphonic audio previously, and 9 of which had experience of critical listening.

- **Reference**
  Every participant was able to identify the hidden reference the vast majority of the time. Only in 5 out of 95 cases was it scored outside of the range of -5 to $+5$. See Figure 6.

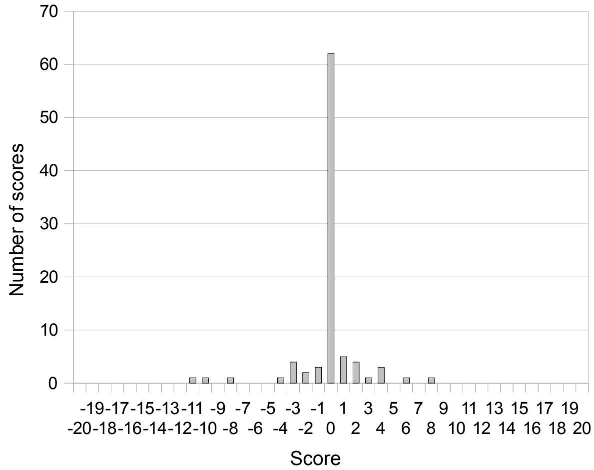

Figure 6: Frequency distribution of the scores given for the ‘Hex’ configuration

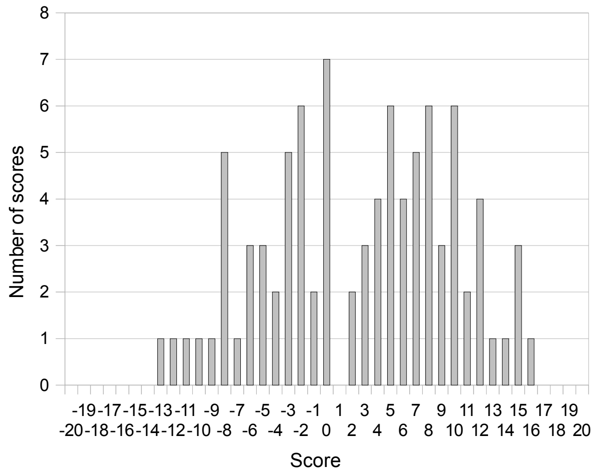

Figure 7: Frequency distribution of the scores given for the ‘HexTri’ configuration

- **HexTri**
  This configuration was expected to perform much better than no height. Although it emerged with the highest overall score, the gap was not as large as expected. Some commented that it didn’t sound very different from the reference, but many more commented on the space, distance and atmosphere that the configuration brought. See figure 7.

- **HexTriNoZ**
  Having lost its height information, this configuration was expected to sound and perform much like the reference. The overall score turned out to be very close to zero, but with a wide distribution of scores above and below zero. Predictably, many commented on its similarity to the reference, or that it sounded ‘flatter’ or ‘duller’. However, some commented on its ‘good height’ or better ‘sense of space’.

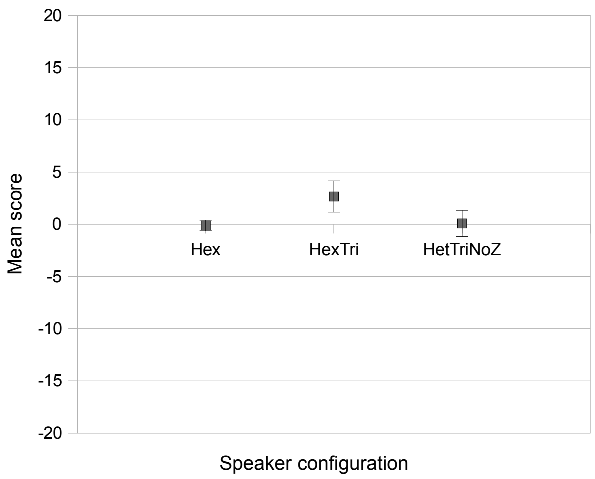

Figure 8: Average scores for each configuration with 95% confidence bars

4.1.5 Conclusions

It is difficult to draw firm conclusions from these results (see figure 8). In the end, the periphonic configuration HexTri emerged with a narrow lead but also drew criticism from many participants. However, it become clear that 'with-height' audio received better feedback in certain situations. There was a clear preference for height when using atmospheric, non-directional content. For items containing sources in-front and in the horizontal plane, such as music, there was no clear preference for the use of height.

Overall it can be said that for some people and situations the effect of height was noticeable and desirable. However, it is not clear why opinion often differed when considering the same content and loudspeaker configuration. The test described here was very general, and did not look at any one aspect of the listening experience. Further work is justified to attempt to better understand how height affects listener experience.

4.2 Comparisons with stereo and 5.0

The vast majority of audio produced by the BBC is in stereo. The notable exceptions to this are certain BBC HD productions and material for DVD release, which are both made in 5.1, and Radio 5 Live which is broadcast in mono.

A listening test was conducted to compare the experience of listening to the same material in stereo, 5.1 and 1st-order Ambisonics. 15 listeners of mixed ability were selected for the panel.

4.2.1 Items

Five items of material typical of that produced by the BBC were selected. Three items were music, two were drama. The nature of the test required separate stereo, 5.0 and Ambisonics mixes for each item, in order to make best creative use of each of the formats. Material which exists in all three formats is in very limited supply, so in most cases mixes had to be created for the purpose of this test. The examples were each 30 seconds in length and are listed below.

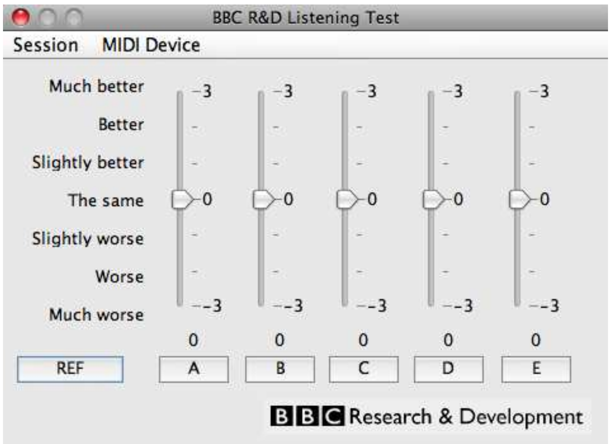

Figure 9: GUI of the second the listening test (the GUI from the first test was almost identical)

1. Classical Music: The BBC Concert Orchestra at the Proms 2009.
   A section from an ensemble piece of classical music played in the Royal Albert Hall.

2. Radio Drama: The Wonderful Wizard of Oz.
   A very dramatic section where Dorothy's house is swept up into a tornado. It features some dialogue and a lot of sound effects.

3. Popular Music: The Get Out Clause.
   A stripped down, intimate acoustic recording.

4. Jazz Music: The BBC Concert Orchestra at the Proms 2009.
   Orchestral Jazz featuring two soloists; a singer and a trumpet player played in the Royal Albert Hall.

5. Radio Drama: The Wonderful Wizard of Oz.
   The opening of the radio play, featuring a voiceover and some dialogue with limited, quiet and spacious sound effects.

4.2.2 Set Up

This test was conducted in a listening room at BBC New Broadcasting House in Manchester using the set-up shown in figure 3. The room contained 16 PMC DB1-SA active monitors, 14 of which were used for Ambisonics playback, with the extra two used for stereo. The 5.0 playback used the two stereo speakers and the front centre, rear left and rear right from the Ambisonic array. The stereo and 5.0 layout followed the recommendations laid out by ITU-R BS.1116-1 [3]. The Ambisonics system used a hexagon with front and rear centre for the horizontal array, and a cube with a square of four speakers above and below to add height as shown in table 2. The listening room was an old radio studio which was acoustically treated with absorptive material to control reverberation times and improve room response.

Speaker	X	Y	Z	Azi	Ele
1	Front	Left	Down	-45°	-45°
2	Front	Right	Down	45°	-45°
3	Back	Right	Down	135°	-45°
4	Back	Left	Down	-135°	-45°
5	Front	Centre	-	0°	0°
6	Front	Right	-	60°	0°
7	Back	Left	-	120°	0°
8	Back	Centre	-	180°	0°
9	Back	Right	-	-120°	0°
10	Front	Left	-	-60°	0°
11	Front	Left	Up	-45°	+45°
12	Front	Right	Up	45°	+45°
13	Back	Right	Up	135°	+45°
14	Back	Left	Up	-135°	+45°

Table 2: Speaker positions used for the playback of the Ambisonic material in the listening test

3	Much better
2	Better
1	Slightly better
0	The same
-1	Slightly worse
-2	Worse
-3	Much worse

Table 3: Preference rating scale for the second listening test

4.2.3 Method

The method of this listening test was based on the MUSHRA standard using the 5.0 signal as a reference. It also used a hidden reference and two hidden anchors. The two hidden anchors were corrupted 5.0 (the rear left and right were replaced with front left and right at -6dB) and corrupted Ambisonics signals (the z channel was removed). Listeners were asked to rate their preference for each example relative to the reference using the scale recommended in Miyasaka [11], as shown in table 3. The items were 30 seconds long but were looped until the subjects were happy to move on. The subjects could switch between stimuli and therefore determine the pace of the test. Figure 9 shows a screen shot of the user interface used by the subjects in the listening tests. In addition to the quantitative results the subjects were asked to provide verbal feedback justifying and commenting on their scores, which is summarised in section 4.2.5.

4.2.4 Quantitative Results

It can be seen from figure 10 that although the Ambisonics is the favourite, when 95% confidence bars are added the distribution of results shows that there is no clear winner when it comes to overall preference. Splitting the results into the two categories, music and drama, gives a little more information. Figure 11 shows that the subjects rated stereo as inferior to both 5.0 and Ambisonics when it came to music. 95% confidence bars show that distribution of preferences

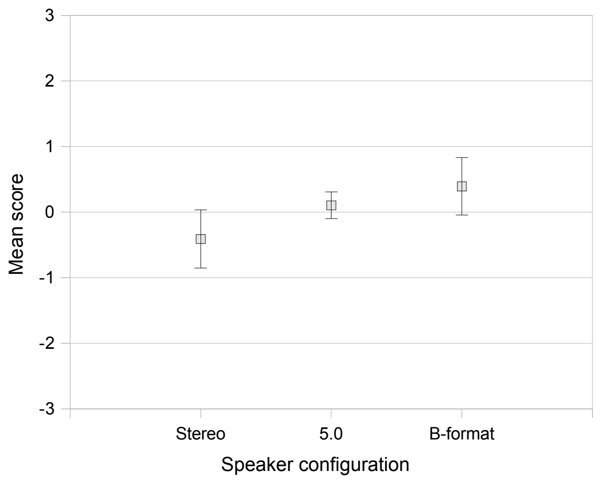

Figure 10: Average preference for all items, shown with 95% confidence bars

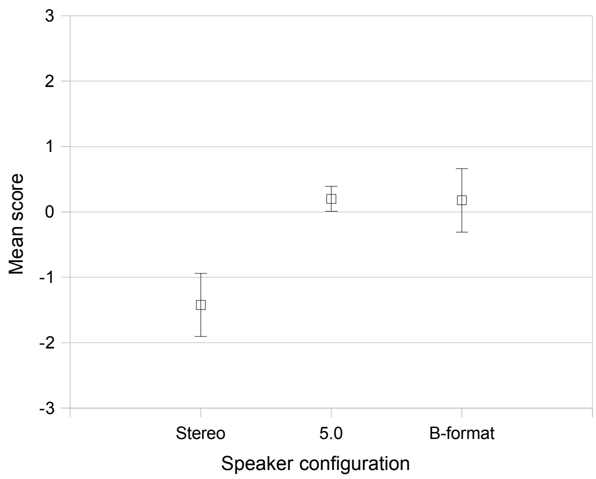

Figure 11: Average preference for musical items, shown with 95% confidence bars

for the Ambisonics was a similar to that of the stereo, unlike the 5.0 which was anchored by the reference.

4.2.5 Qualitative Results

Despite the fact the quantitate results didn't show a clear preference, qualitative results will be presented to further explain the data. The stereo mix tended to receive negative comments from subjects when compared to the reference. 35% of the comments about stereo were very negative, the classical and jazz music receiving the majority of the negative comments. For the two drama examples more than half of the subjects tested said they struggled to grade the Ambisonics because they preferred the Ambisonic representation of the sound effects but not the dialogue. Comments such as "the sound effects are really good but vocal is not so good, she sounds muted, the sound effects are lovely" were typical. 52% of the comments about Ambisonics said that the subjects felt like they were present at the performance, but interestingly this did not always correlate with them preferring the experience. Comments such as "I feel like I'm too in it" and "You feel like you're in it, but not listening to it" were typical.

4.2.6 Conclusions

The quantitative results are somewhat inconclusive, but when combined with qualitative results tentative conclusions can be drawn. The preference for Ambisonics seemed to show some correlation with the type of material. A preference for the Ambisonics stimulus was shown with ensemble music featuring considerable spatial effects and sound effects which surrounded the listener. A preference for 5.0 was shown when the clips featured more obvious point sources, such as dialogue which included limited room effects or narration which featured no room effects.

It would be of great interest to discover whether higher order Ambisonics or a hybrid method of presentation could address these issues.

5 Barriers to Adoption of Ambisonics

Interest in Ambisonics outside the academic and enthusiast communities appears to have risen in recent years. This is most notable in computer games publishing, where gamers have been early adopters of surround sound, and the nature of computer games means that complex DSP can be used for audio. There remains a number of barriers to the adoption of Ambisonics in broadcast production workflows, and the case studies presented in section 3 identify a number of them.

5.1 Production Tools

Perhaps the most noticeable barrier to the adoption of Ambisonics in production is the lack of suitable hardware and software tools. There is a large number of tools available for the coding, manipulation and decoding of Ambisonic audio, many of which have been developed by the academic and enthusiast community for their own purposes. However, most of these tools are unsuitable for use in a broadcast environment. Their documentation tends to be sparse, and it is often unclear exactly what processes these tools are applying to the audio. These tools are often VST (Virtual Studio Technology) plug-ins and are limited to one system platform. This can cause problems in many broadcasting environments, where a wide variety of different operating systems and editing packages may be in use. The majority of these tools currently only handle up to first order Ambisonics, not allowing integration of higher order Ambisonics into workflows. In addition, most audio recording software and digital desks are designed for discrete channel systems such as stereo, 5.1 and 7.1, and the routing of non-standard groups of channels is not trivial for non-technical users.

5.2 Decoders

Decoders are also an issue. Listening rooms and speaker set-ups vary considerably, making Ambisonic decoding a complex process. A number of decoding software packages are available, but most require the user to chose a preset speaker layout, or create their own decoding matrix. Interoperability is generally limited with decoders either running as VST, or requiring other virtual routing tools such as JACK. A better solution would provide the user with an easy-to-use interface to specify speaker locations, and support features such as higher order decoding, different weighting options, shelf filtering and near field compensation.

5.3 Microphone Choice

The recordings described in this paper used sound field microphones, because they were the only type that were easily available. Engineers tend to have a preference for different microphones in different areas and like to be offered a choice. Higher-order multi-capsule microphones such as the MH Acoustics “Eigenmike” [12], are in development, but for the moment are expensive and not readily available.

5.4 File Format

There is also a lack of standardised file formats for Ambisonics. Whilst there are channel-ordering conventions for B-format and higher orders, the increasing dominance of file-based production techniques means that standardised file formats and metadata are essential for effective broadcasting and archiving.

5.5 More Demonstrations

Finally, there seems to be a lack of convincing Ambisonics demonstrations. It could be argued that one of the reasons Ambisonics has failed to move into the mainstream is little or bad marketing. Ambisonics needs to prove itself though a greater number of excellent demonstrations in order to convince more influential industry figures.

6 Conclusions

Content producers are excited by the creative opportunities that Ambisonics presents, but there is a shortage of tools available that are suitable for broadcasting.

Subjective tests showed that first-order Ambisonics seems to offer advantages over 5.1 with certain types of material – namely sound effects and ambience – but that it suffers with dialogue and narration. Similarly, the height dimension works better when replaying content where the sound sources do not lie on the horizontal plane, but are spread across the sound field.

7 Future Work

This paper answers a number of important questions, but it also raises many. Although the quantitative results are generally inconclusive the qualitative results suggest future areas of research. The most obvious step would be to conduct similar case studies and listening tests using higher order Ambisonics. There could also be an investigation of hybrid methods of delivery which could use some combination of Ambisonics and discrete channel systems. An investigation into how the lossy audio compression technologies employed in the broadcast chain would affect Ambisonics signals would also be beneficial. It would also be informative to conduct subjective testing in a ‘normal’ domestic environment with a relatively low cost Ambisonics set-up, to simulate a typical audience member’s listening environment.

8 Acknowledgments

The authors would like to extend their thanks to Simon Tuff, Rupert Brun, Steve Brooke, Nadia Molinari, SIS LIVE, Richard Furse, Simon Goodwin, Bruce Wiggins and Peter Lennox for their help and input. Special thanks go to Andrew Mason and David Marston for their invaluable knowledge and assistance.

References

[1] BBC website, "Public purposes," http://www.bbc.co.uk/aboutthebbc/purpose/

[2] M. Neukom, "Ambisonic Panning," AES Paper, 2007 October.

[3] ITU-R recommendation BS.1116-1, "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems," 1997 October.

[4] BBC website, "Proms 2009," http://www.bbc.co.uk/proms/2009/

[5] BBC website, "The Wonderful Wizard of Oz," http://www.bbc.co.uk/programmes/b00pb8x1

[6] B Wiggins, "An Investigation into the Real-Time Manipulation and Control of Three-Dimensional Sound Fields," University of Derby, 2004.

[7] Hamasaki, K; Hiyama, K, "Reproducing Spatial Impression With Multichannel Audio," AES Paper, 2003 June.

[8] Soundfield website, "Soundfield DSF-2 Digital Broadcast Microphone System," http://www.soundfield.com/products/dsf2.php

[9] Wikipedia, "Foley artist," http://en.wikipedia.org/wiki/Foley_artist

[10] ITU-R recommendation BS.1534. "Method for the subjective assessment of intermediate quality levels of coding systems (MUSHRA)," 2003 January.

[11] E. Miyasaka, "Methods of Quality Assessment of Multichannel Sound Systems," AES Paper, 1993 June.

[12] MH Acoutics Website, "Eigenmike," http://www.mhacoustics.com/