

# Ambisonic Sound Design for Theatre with Virtual Reality Demonstration - A Case Study

Alexander Vilkaitis<sup>1</sup> and Bruce Wiggins<sup>2</sup>

<sup>1</sup> University of Derby, Derby, UK  
a.vilkaitis1@derby.ac.uk

<sup>2</sup> University of Derby, Derby, UK  
b.j.wiggins@derby.ac.uk

## Abstract

This paper discusses ambisonic sound design for a theatrical production of King Lear. Sound, and its use in theatre, has taken a back-seat in comparison to the development of other theatre technologies such as lighting, projection and automation in recent years. Spatial audio implementations in theatre give the sound designer and the artistic team much greater scope for creativity, along with improvements in source separation and intelligibility of sources due to spatial unmasking. A 360 degree video was also recorded, with first and third order ambisonic binaural reproductions of the sound design stitched on to the video to create a virtual reality experience. The project was successful, whilst highlighting some practical and perceptual limitations in spatial audio for theatre.

## 1 Introduction

The development and use of sound in theatre has taken a relative back-seat compared to other technological aspects of productions. In recent years, the use of lighting, automation and projection have all been used in new and innovative ways in landmark productions such as the projection in the National Theatre’s “Curious Incident of the Dog in the Night Time” [15], the automation and rigging in “The Light Princess” [16] and the projection mapping in the Royal Opera’s “Don Giovanni” [3]. However, the use and implementation of sound has been relatively static in comparison. The recent increase in enthusiasm due surround audio for Virtual Reality; mostly utilising ambisonics, need not only take place in the virtual world, or in the cinema. It is time the gap is bridged, and theatre as a form of virtual reality itself, takes advantage of immersive audio.

Current implementations of theatre sound design are often only as reinforcement tools for the performers and band/orchestra, with occasional ambient effect playback. Until recently, theatrical productions have been reproduced in stereo, with either mono reinforcement, or basic stereo panning dependent on source location. The use of spatial audio techniques, to spread both reinforced and playback sounds not only across the stage area, but around the auditorium is surely a vast improvement over stereo systems, allowing audio to break the fourth wall and

immerse the audience.

There is scope for surround sound implementations in theatre as localisation and movement of sound sources improves realism and immersion in theatrical productions [7]. There is not only an improvement in experience available, but greater creative scope for the artistic team, as well as quantifiable benefits in terms of intelligibility, source separation, and a reduction in the need to heavily spectrally equalise in order to have all sound sources separate in the mix [9].

This research documents the process of creating and implementing an ambisonic sound design for a theatrical production of Shakespeare’s *King Lear*. The sound design process begins with a dialogue with the director, and adjust to his or her concept or setting of the production. An ambisonic based surround design was chosen as the play was set in a 1960’s catholic mental asylum, giving fantastic scope for creative sound design, and to immerse the audience in the mis-en-scene of the play. Ambisonics is particularly suited to creating a sense of ambience, as the audience are situated within the sound field.

The venue for the production was a grade II listed building that has been converted into a theatre and ran by Derby Live. It has a capacity of circa 200 people including the balcony, though the balcony was not used as an audience area but as a technical operating point for this production.

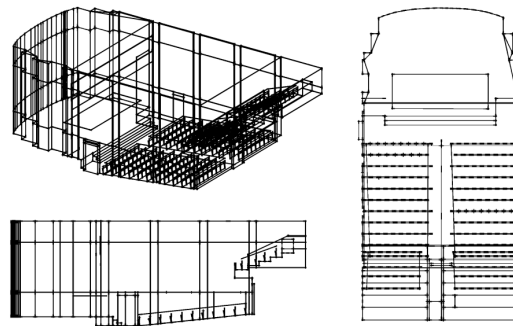


Figure 1: Plan, profile and isometric view of a model of The Guildhall, Derby, UK. [6]

## 2 Literature

### 2.1 Spatial Audio

There are two main methods of spatial content creation, scene based and object based; both of which have advantages and disadvantages. Scene based audio has relative advantages in that the rendering complexity is much lower compared to that of object based, as well as providing a greater spatial fidelity and resolution than object based systems. The encoding and decoding of scene based audio are also separate, allowing for one mix to be rendered for any speaker configuration, or even for binaural reproduction over headphones [14].

Object based systems, like scene based, are format agnostic but work on the premise of separating sound files into audio objects with associated metadata describing various traits

such as position and level, which are then rendered in real time depending on the speaker layouts. Object based audio is much more computationally taxing as it requires real time computation [1].

In a practical theatrical implementation, either method has advantages. Scene based audio lends itself to the structure of traditional theatre sound workflows, which are already scene based, however on-the-fly editing would not be possible as the soundscapes are pre-rendered. Object based audio implementations however would allow real time manipulation, allowing the sound engineer to adapt to any on stage mishaps.

Spatial unmasking is used in loud or complex auditory environments, and is known as the “cocktail party effect” which is achieved by separating out auditory events in space. The cocktail party effect describes the ability for the auditory system’s ability to discriminate between sound sources at different locations in space [10]. Using spatial reproduction, and therefore spatial unmasking, it allows the sound designer to have less need for equalisation and dynamics processing, as source separation in the spectral domain is no longer the sole method of discriminating auditory events. The ability to place sounds throughout the auditorium also decreases the masking taking place when sound is only sourced at the proscenium arch, near-coincident with the actors. In turn, this separation of sources and subsequent spatial unmasking will increase intelligibility.

## 2.2 Ambisonics

Ambisonics is a method of spatial reproduction that aims to reproduce the original sound field based on spherical harmonics in both azimuth and elevation planes. Spherical harmonics represent functions on a sphere, in this case sound pressure levels [12]. The minimum number of channels required for full 3D spatial reproduction is four, W, X, Y and Z which represents polar patterns of an omnidirectional and three figure of eights; known as B-format. Increasing the number of channels, and therefore the *order* increases the spatial resolution (Fig. 2).

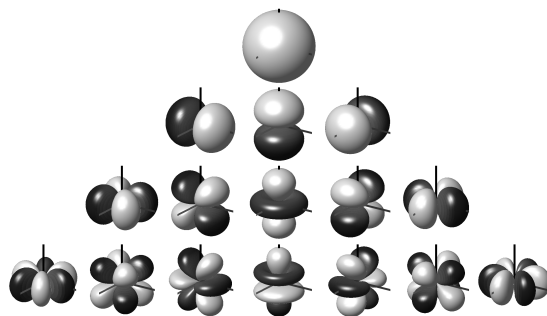


Figure 2: Spherical harmonics for 0-3rd order ambisonics [19]

The benefit of ambisonic reproduction as opposed to traditional stereo is that there is no longer separation between the proscenium arch and the audience, as they share the world created on stage by being surrounded by sound.

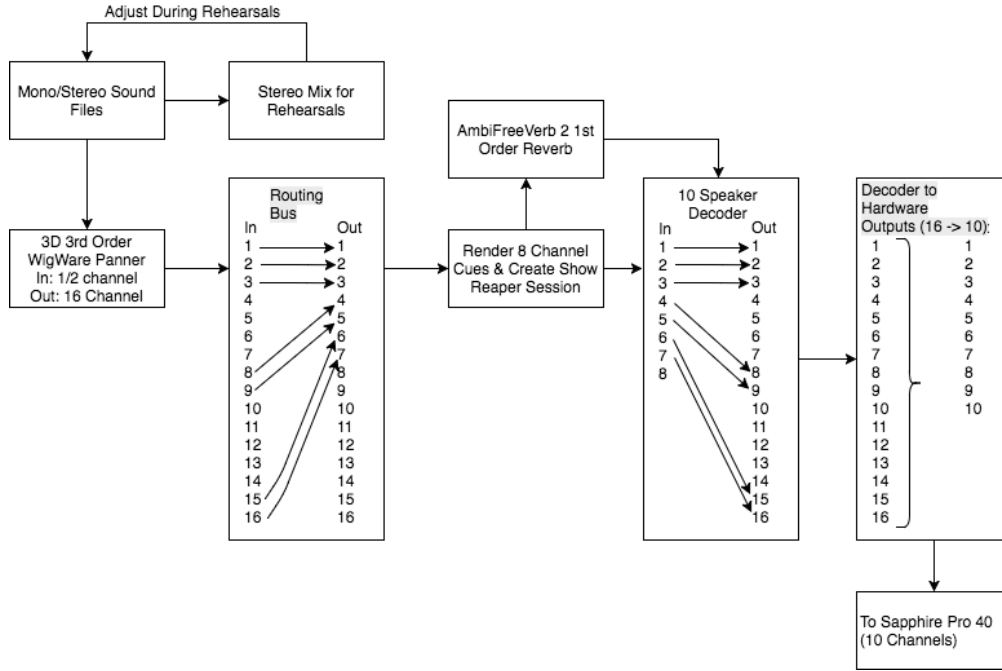


Figure 3: Reaper project diagram with routing

### 3 Method - Ambisonic Sound Design for Theatre

Traditional Higher Order Ambisonics (HOA) uses large speaker arrays to reproduce a given 3D sound field. For live events, full 3D reproduction at higher order is impractical for the scope of this project due to the large number of speakers required, therefore a 2D solution was used. The minimum number of channels required to reproduce a given ambisonic order in 2D is given by:

$$N_{speakers} = 2n + 2, \text{ Where } n = \text{order} \quad (1)$$

The design process began with the creation of stereo mixes of each sound cue as required by the director, as a full surround array was not available for use in rehearsals. Mono sound files from FreeSound [4] were layered, edited and panned using 2D third order ambisonic panners (7 channel) [18] in Reaper [13], before being summed to a 16 channel bus, and decoded using a 3D 10 speaker regular 3rd order decoder (Fig: 3). The sound design was finalised in a studio dedicated to ambisonics with a 32.2 3D speaker array, which was then adapted to a 2D 10 speaker setup by changing channel routing and the decoder, which highlights the flexibility of ambisonics as a system that allows the preparation work to take place in one studio before being transferred to the venue which would reduce venue hire costs significantly.

This implementation required 2D reproduction, and therefore the system can be optimised by disregarding the channels containing elevation information. This reduces the 16 channels of 3rd order to only 7 channels. Using (Eq. 1), the minimum number of channels for 3rd order 2D reproduction is 8, however due to the nature of the venue, a 10 channel system was utilised.

Each individual 3rd order sound cue was placed into a Reaper project file with the required routing, and marked with markers. This enabled the sound operator to trigger each cue as

required by using a number hotkey. This is a departure from traditional theatre sound control software such as QLab [8], where cues are placed in lists. However, QLab does not support multichannel spatial audio so Reaper had to be used due to its highly customisable routing system and support for ambisonics.

A Sapphire Pro 40 was used as the interface between a Macbook Pro and the 10 individual “flying mole” amplifiers [2] which outputs to 10 Podspeakers [11]. Originally, the speakers were placed above head height on tripod speaker stands, however these were too intrusive in relation to the auditorium seating positions, and from a visual perspective and as a result the stands were removed and the speakers were placed on the floor. This highlights the need for any surround configuration in theatre to be as visually discrete as possible.

The speaker array consisted of a left and right pair either side of the proscenium arch, with two down each side of the auditorium, one in each rear corner, and two at the rear mirroring the frontal pair. While the array shape was far from ideal in-terms of spherical reproduction accuracy, compromise is needed between spatial accuracy and intrusiveness of the system. The sound cues were designed with these compromises in mind as low spatial resolution was required around the front third of the array with greater resolution towards the rear which in turn suited the practical implementation due to the long and narrow shape of the venue.

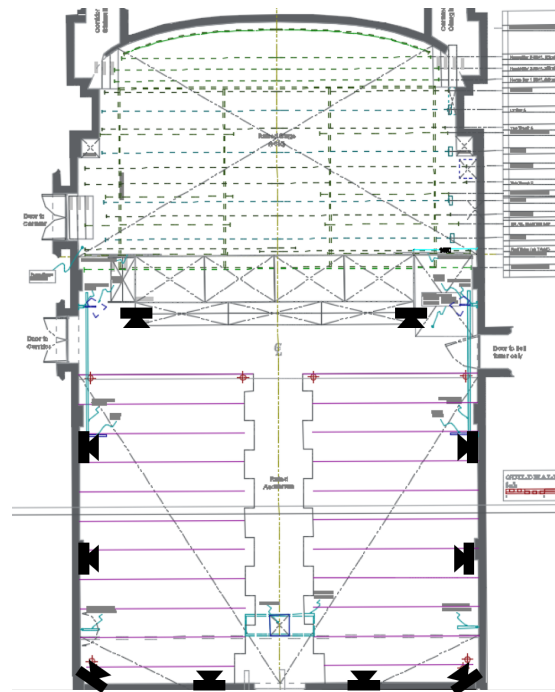


Figure 4: Speaker positions in the Guildhall Stalls

## 4 Results

### 4.1 Practical Considerations

There are many practical considerations when implementing a spatial audio sound design for theatre. The theatrical context stipulates that any sound system be as discreet and unobtrusive as possible whilst providing good intelligibility, even Sound Pressure Level (SPL) over the entire audience area, and be as low cost as possible as sound reinforcement is only a small part of a technical budget covering lighting, rigging, set manufacture and stage crew.

Another practical issue is the amount of cabling required to implement even a 10 channel ambisonic rig. As each speaker is on a separate channel, daisy chaining is not possible and each speaker requires a separate NL4 cable from the amplifiers to the speakers distributed all around the auditorium. This leads to a large distribution of cables which is difficult to manage in a tidy and orderly fashion.

The monitoring/operating location is also critically important. The lighting and sound operating area was placed at the back of the upper balcony, which makes monitoring of both sound levels but also spatial quality impossible as the operator is outside the ambisonic array. A possible solution would be a more elaborate setup with a digital mixing desk that allows control via an iPad OSC network client which would allow the sound engineer to sit in the auditorium during technical rehearsals. Another option would be monitoring via a live feed binauralisation of a spatial microphone placed in the auditorium.

### 4.2 Perceptual Evaluation

The sound design was successful, with cast members, audience and the directorial team being impressed with the implementation and creative possibilities it can provide. There are perceptual issues regarding audience position and their individual experience due to loudspeaker proximity and auditory masking of un-amplified dialogue creating an different listening experience throughout the auditorium.

This ability to spread the sound around the auditorium also leads to spatial unmasking of sources, as the sound effects are no longer sourced concurrently at a standard left and right stereo based system. This increases the intelligibility of the un-amplified voices for most of the audience as they are not masked by sound effects coming exclusively from the proscenium arch.

The proximity of the speakers to the audience was an issue, as several audience members were seated very close to individual speakers, meaning that the actors on stage were masked occasionally by the ambient sound design. Placing the speakers on the floor of the auditorium helped reduce this masking as the speakers were no longer at head height, however the issue persisted. This could be reduced by placing the speakers further from the audience in a larger venue, essentially creating a central audience area in the *sweet spot* with each speaker some distance away from the audience, however this was not possible in the relatively small venue. This approach would also have an impact on the number of seats available for ticket sales. Another option would be to have a greater number of speakers at a lower SPL, or by heavily diffusing speakers by placing an obstruction directly in front of them.

Due only having two pod speakers at the front of the auditorium, the front image was not as strong and present as it should be. A possible solution would be to increase the number of speakers around the front, creating a front weighted speaker array, or use traditional PA speakers at the proscenium arch that are capable of producing greater sound pressure levels. This would skew the ambisonic image to be more like a traditional stereo implementation, however the sense of ambience from the surround speakers would pervade. A more elaborate

solution to give greater spatial resolution at the front could be implemented by using a planar wave field synthesis array and distributed surround speakers, however this is a topic of future research.

## 5 Virtual Reality Demonstration

A section of the play without dialogue was recorded in 360° using a Ricoh Theta S camera. First and third order renderings of the sound design were then stitched on to the 360° video using Google Spatial Media Tools [5]. This forms a basis of a head tracked VR demo. A previous paper by the authors explores VR spatial audio video creation in greater detail [17].

## 6 Conclusion

Spatial sound design for theatre opens a plethora of creative opportunities which create an immersive experience for the audience. The breaking of the fourth wall and involving the audience as part of the *mis-en-scene* removes the barrier between the world of the stage and the auditorium. There are practical limitations, and further work to improve the perceptual experience across the entire performance space needs to be carried out. The efficacy of VR as a demonstration platform for spatial audio and live theatre also needs to be investigated.

## References

- [1] Philip Coleman, Andreas Franck, and et al. On object-based audio with reverberation. *Audio Engineering Society*, 2016.
- [2] Flying Mole Corp. Flying mole amplifiers. [https://www.flyingmole.co.jp/new\\_en/products/new\\_product/dadm100pro2/](https://www.flyingmole.co.jp/new_en/products/new_product/dadm100pro2/), 2017.
- [3] Emma Baker and (ROH Magazine Editor). Trade secrets: How do you 'paint' a rotating set with video projections?, January 2015.
- [4] FreeSound.org. Freesound. <http://freesound.org>, 12 2016.
- [5] Google, Inc. Using the spatial media metadata injector. <https://support.google.com/jump/answer/7044297?hl=en>, 2016.
- [6] Pete Jones. The guildhall derby, a sound system design and room acoustic analysis. April 2017.
- [7] D. Kaye and J. Lebrecht. *Sound and Music for the Theatre: The Art Technique of Design*. Focal Press, 2016.
- [8] Figure 53 LLC. Qlab. <https://figure53.com/qlab/>, 2017.
- [9] Julian Palacino, Mathieu Paquier, Vincent Koehl, Frédéric Cjangenet, and Etienne Corteel. Assessment of the impact of spatial audiovisual coherence on source unmasking - preliminary discrimination test. *Audio Engineering Society Convention 140*, June 2016.
- [10] Georg Plenge and U Tilse. The cocktail party effect with and without conflicting visual cues. *Audio Engineering Society Convention 50*, 1975.
- [11] Podspeakers. <https://www.podspeakers.com/>, Title = Podspeakers, Year = 2017.
- [12] Boaz Rafaely. *Fundamentals of Spherical Array Processing*. Springer Verlag, 2015.
- [13] Reaper. Reaper. <http://www.reaper.fm/>, 2017.
- [14] et al Shankar Shivappa, Martin Morrell. Efficient, compelling and immersive vr audio experience using scene based audio/higher order ambisonics. *Audio Engineering Society*, 2016.
- [15] The National Theatre. Curious incident of the dog in the nighttime.

- [16] The National Theatre. The light princess background pack - interview with paul rubin, aerial effects designer. Online, 2013.
- [17] A. Vilkaitis, M. Dring, C. Middlicott, B. Wiggins, and A. Hill. Room acoustics and virtual reality: An implementation of auralisation and 360 degree image techniques to create virtual representations of spaces. *Reproduced Sound 2016*, November 2016.
- [18] Bruce Wiggins. Wigware. [http://www.brucewiggins.co.uk/?page\\_id=78](http://www.brucewiggins.co.uk/?page_id=78), 2016.
- [19] Franz Zotter. Spherical harmonics up to degree 3, as used in third-order ambisonics. [https://commons.wikimedia.org/wiki/File:Spherical\\_Harmonics\\_deg3.png](https://commons.wikimedia.org/wiki/File:Spherical_Harmonics_deg3.png), 2013.