THE DESIGN OF AN *AUDIO FILM* FOR THE VISUALLY IMPAIRED

Mariana Julieta Lopez and Sandra Pauletto

Department of Theatre, Film and Television
The University of York, Heslington, York, YO10 5DQ
marianajlopez@yahoo.com, sp148@york.ac.uk

1. ABSTRACT

Nowadays, Audio Description is used to enable visually impaired people to access films. However, it presents an important limitation, which consists in the need of the visually impaired audiences to rely on a describer, not being able to access the work directly.

The aim of this project was to design a format of sonic art called *audio film* that eliminates the need of visual elements and of a describer, by providing information solely through sound, sound processing and spatialization, and which might be considered as an alternative to Audio Description.

In order to explore the viability of this format an example has been designed based on Roald Dahl's *Lamb to the Slaughter* (1954) using a 6.1 surround sound configuration. Through the design of this example it could be noticed that this format can successfully convey a story without the need of either visual elements or of a narrator.

2. INTRODUCTION

Film, television, theatre performances and museum tours use Audio Description to enable visually impaired people to access these forms of art. However, the inclusion of these descriptions has as a consequence that visually impaired audiences cannot access the work directly but have to rely on a describer.

The aim of this project was to design an alternative to Audio Description for films.

In order to explore the potential of this format an example based on Roald Dahl's *Lamb to the Slaughter* (1954) was designed for a 6.1 surround sound setup.

Lamb to the Slaughter is the story of Mary Maloney, who after finding out that her husband plans to leave her, murders him with a frozen leg of lamb and feeds the murder weapon to the police.

In this project the term *audio film* was chosen due to the belief that certain elements from the filmmaking process might be adapted for the conveyance of a story through sound, creating an experience equivalent to the cinematic experience.

In this project the adaptation of different cinema languages was explored, in particular the use of the Master Scene and the interpersonal cinema language.

Ron Richards defines the Master Scene as the use of one shot 'that establishes the environment and the people, and records the event or action in its entirety.' [1, p.89] This can be used to introduce the different spaces by presenting all the sounds heard in that particular environment. In this way the space is aurally established as it would be visually established through a long shot.

Also, the interpersonal cinema language, which focuses on involving the audience with the emotional states of the characters, can be achieved through the content of the lines delivered, the expressiveness in the voices and the use of music to emphasize the characters' feelings.

The notion of cuts through sound can be explored to indicate parallel actions taking place in different spaces.

An effect comparable to that of a tracking shot can be achieved by editing the sounds in a way in which a character's movements are followed through different spaces.

2.1. Audio Description

The design of an *audio film* is motivated by the possibility of providing an alternative to Audio Description.

The concept of Audio Description was first introduced and studied by Gregory Frazier in 1975 and consists in an audio track in which

'a describer inserts spoken words to provide representations of information contained in the visual field of the production' [2, p.1].

An important analysis on Audio Description is the one provided by Philip J. Piety. Piety points out four components of Audio Description: *source text*, *modified text*, *describer* and *consumer* [2, p.28]. The *source text* is the original text, generally aimed at a sighted audience. The *modified text* is the conjunction of the source text and the descriptions that have been inserted. The *describer* refers to the person in charge of the creation and the placement of the descriptions and the *consumer* to the recipient of the modified text [2, p.29].

Although Audio Description has enabled visually impaired people to access visual forms of art it presents several limitations.

Firstly, it evaluates and summarizes the scene described, leaving the consumer 'less opportunity to independently assign meaning' [2, p.35].

Secondly, Audio Description must not overlap the dialogues. This means that visual aspects are not described during dialogues, leaving the listener without part of the information. Another problem regarding this is that descriptions are constrained to the gaps between lines, which may not be long enough to provide the necessary descriptions.

A further limitation is related to the situations in which descriptions are synchronized to the sound effects. Descriptions mask what might be important audio clues, which might be interpreted without the need of descriptions.

Another limitation is that consumers cannot experience the source text on its own because it does not contain enough information for them to follow the plot.

The *audio film* proposed attempts to solve most of these limitations by providing just a *source text* that enables the visually impaired listener to experience the *audio film* in the same way as it would be experienced by a sighted listener. Also, by eliminating descriptions, the listener will be able to use his/her imagination more freely. Furthermore, the fact that this format does not need a describer, would eliminate another disadvantage of Audio Description: the way in which the

describer decides to present the information modifies the way in which the film is experienced [2, p.73].

2.2. Audio-Only Games

The design of an audio film can learn from techniques used in Audio-Only Games.

An interesting approach is the one taken by the TiM (Tactile Interactive Multimedia) project carried out by SITREC (Stockholm International Toy Research Centre).[1]

SITREC has worked on the design of *Tim's Journey*, an adventure game based on surround sound, which takes place on an island divided into different spaces. In order to enable the player to recognize them, each of these spaces is characterized by a musical theme. Also, to make navigation easier the footsteps of the avatar provide information regarding the surface it is walking on [3].

Both design strategies will be used for the design of an *audio film* to clarify the plot.

2.3. Radio Drama

The concept of *audio film* bears some resemblance to radio drama. The main differences being that an *audio film* does not use narration and employs surround sound to convey information.

William Ash defines a radio play as 'a story told in dramatic form by means of sound alone.' [4, p.1] This definition points out the main aspect of a radio play: its sole reliance on sound for storytelling.

Rattigan asserts that radio drama is a format in itself that does not need visual elements, it is

'not handicapped by the absence of any visual output, on the contrary, its 'sightlessness' is the basis of its unique appeal, which promotes an imaginative visualization on the part of its listeners' [5, p.1].

An *audio film* is also a format in itself that does not need images. On the contrary, it is designed with a purely aural approach, which stimulates the listener's imagination.

An audio film also includes the same elements present in the radio drama: silence, pauses, voice, sound effects, utterances and music [5].

As regards sound effects, Rattigan argues that they need a verbal contextualization to acquire meaning for the listener regarding their intention and their consequence. At the same time he warns that the intentions stated should not be naïve or overstated [5, p.154].

Crisell also suggests this need of contextualization by stating that 'it seems doubtful whether any radio sound is ultimately meaningful without the help of speech' [6, p.141].

Although it could be agreed that sound effects, which are not contextualized, might disorient the audience, speech contextualization might not be the only alternative. Music could be used to denote certain meanings in those cases in which verbal statements would destroy the mood of a scene of a radio drama or an *audio film*.

Furthermore, it seems that by considering that sound effects must be contextualized through speech, a great stress is being placed on the voice while denying the possibility of communicating full meaning through non-speech sounds.

These opinions seem to underestimate the ability of the listeners to put together different sound effects in their minds to

reconstruct meanings or to use other clues other than speech to interpret those sounds. It also seems to be suggested that sounds need to be clarified immediately when it might occur that meanings become clearer as the story evolves in time and in the listeners' minds.

In the design of the *audio film* the different uses of music in films are also of relevance, including its use to convey information about a character's emotional state, create a general atmosphere or a specific mood, and its employment as a means of narration [7].

3. THE PRODUCTION STAGE

3.1. The Adaptation Of Roald Dahl's Story

A script was written providing details on the sound elements that conform the soundscape as well as the actions included. Details on the feelings of the main characters were included both to provide guidance for the performance and to be used as hints for the selection of music. Dialogues and internal monologues were based on Dahl's story as well as on Hitchcock's adaptation for the series *Alfred Hitchcock presents* (1958).

3.2. Recordings and editing

All voice recordings were done using an AKG C 414 B-XLS microphone. In addition to the dialogues, the actors were asked to perform breathing and chewing sounds and utterances. In the case of Mary Maloney, different recordings of sobbing and laughs were also done.

The sound effects employed in *Lamb to the Slaughter* include original recordings (done in studio and on location) as well as sounds taken from sound libraries (*Sound Ideas Series 6000 General Sound Effects Library*, http://www.soundsnap.com/ and *Hollywood Edge Sound Effects Library*).

The editing process was done using Logic Pro 8, with exception of the blow with the frozen leg of lamb, which was designed using Pro Tools LE 7.4.

3.3. Dialogues

It was decided that some of the dialogues ought to be superposed to achieve a more dynamic and realistic work. This choice was based on Robert Altman's concepts on film sound. In order to achieve life-like movies, Altman concentrated in his work

'not in rapid alternating of short lines of dialogue but on dialogue simultaneously spoken by two or more characters: a wall of sound, a Tower of Babel' [8, p.349].

3.4. Soundmarks

Sonnenschein defines a soundmark as a type of sound that 'establishes a particular place, as does a landmark, possessing some unique quality for only that location' [9, p.183].

Soundmarks have been used for indoor and outdoor spaces to help the listener in their recognition. Indoor soundmarks include a cuckoo clock and a freezer while outdoor soundmarks include two types. The first conveys the idea of a residential

---

[1] SITREC is a research centre within KTH, Royal Institute of Technology.

area during daytime while the second is used to communicate the idea of night-time in the same area.

3.5. Footsteps

Footsteps were considered of the utmost importance since they are the means by which movement through different spaces is represented. When editing the footsteps different issues were considered. Firstly, they were deemed essential as a means to communicate information regarding the movements of the characters as well as their emotional states.

Secondly, an attempt was made to create a sense of distance through the footsteps. The number of footsteps that it takes a character to get from one space to another should be as consistent as possible.

Thirdly, footsteps are used to remind the listeners of the presence of those characters that in certain moments might not be taking part in dialogues but are meant to reappear later on in other conversations.

3.6. Character-Object Interaction

This category refers to those sounds that represent objects with which the characters interact. When necessary careful attention was paid to present complete sequences of actions. However, in other circumstances it was considered that reducing the number of sounds used to represent an event was more effective.

On the other hand, some sound effects were included because they clarified the meaning of certain events. This is the case of the door bell included when Mary enters the grocery.

Sound layering was also used to attempt to communicate meaning more effectively. This technique was used in the design of the murder scene in particular for the sound of Mary hitting Patrick with the frozen leg of lamb.

It is essential to mention that the character-object interaction sounds are the ones listeners might have more difficulty in assigning to a particular *landscape*, this is, 'the source from which we *imagine* the sounds to come' [10, p.136].

However, in the editing process it was considered that contextualization might aid the listener in the recognition of sound sources. This contextualization is determined by other sound effects as well as by dialogues, music, and spatialization.

3.7. Internal Sounds – Utterances

In the *audio film* Mary was assigned internal sounds [11] in the form of breathing to try to generate sympathy for her character and aid the listener in the understanding of her state of mind. When no dialogues are present verbal exclamations were also used to make actions clearer to the listeners and avoid confusion.

Utterances were also important to indicate that although some of the characters might not be delivering lines they are present in the different scenes.

3.8. Music

Two music pieces were included in *Lamb to the Slaughter*: Steve Reich's *New York Counterpoint* (Second Movement) and Toru Takemitsu's *Gitimalya*.

New York Counterpoint was used at the beginning to convey the idea of tranquillity and expectation described in the script. It was chosen as a means to indicate Mary's initial

sweetness towards her husband. It was decided that the theme would be used to accompany Mary's attempts to communicate with Patrick, and the music would fade out as those attempts fail.

Gitimalya is used throughout the *audio film* with different functions. Firstly, music is employed to clarify Mary's feelings towards the different situations she goes through. Secondly, music is also used as a tool to create tension. This is particularly important in the murder scene in which music was employed to enhance the tension and impact the listeners.

Thirdly, the marimba theme present in Takemitsu's work is used as a *leitmotif*, this is, a musical theme assigned to a 'main character or key thematic idea of the narrative' [11, p.51]. This theme is used as a *leitmotif* for the murder weapon. It is included every time the subject of the weapon is mentioned and it was included to aid the listener in the understanding of the story. If the listener does not grasp the fact that Mary killed Patrick with the frozen leg of lamb this might help him/her understand this key idea.

4. SOUND PROCESSING

4.1. Voices

An important aspect of sound processing was the use of reverberation to provide the listener with information about the different aural spaces and provide every sound with an environmental context [12] that might allow the listener to identify where an event is taking place. This is of particular importance in the case of visually impaired people since they have a strong motivation to develop spatial awareness due to its importance as an orientation tool [13]. For this reason, artificial reverberation was considered essential to enable listeners to recognize the different spaces presented.

The voices were processed using the Space Designer plug-in of Logic Pro 8, and a different setting was assigned to each space.

Reverberation was also applied to indicate that some of the lines delivered by Mary were part of her thoughts.

4.2. Footsteps

Footsteps were considered an important element since they are a way in which listeners may be aware of certain characters and might even identify them. Consequently, the footsteps belonging to different characters needed to be differentiated through sound processing to try to aid the listener in their distinction. In order to endeavour to achieve these differences the Channel EQ plug-in of Logic Pro 8 was employed.

Footsteps were also processed with the Space Designer plug-in to convey the idea of the different surfaces characters walk on. Furthermore, the same reverb settings used for the voices in the different spaces were mostly maintained.

4.3. Sounds Heard Through Windows

Another important part of the sound processing stage was the processing of those sounds that were supposed to be heard through windows. To attempt to create this effect reverb was added by means of the Space Designer plug-in and equalization was applied to reduce the low frequencies and boost the high frequencies.

4.4. Sounds Heard From Other Rooms

Another effect added through processing was that of sounds being heard from other rooms. This effect was achieved by employing the Space Designer plug-in. This effect was of particular importance at the end of the *audio film*. It was a creative decision to end the *audio film* by simulating a cut from the kitchen where the policemen are having the leg of lamb and commenting on the murder weapon to Mary listening to their conversation in the living room. This is of significance since until that moment all the *audio film* had been centred in the spaces in which Mary was, and this is the only moment in which the listener is taken to a room Mary is not in. To emphasize this the sound of the clock ticking, which characterizes the living room, was processed with this reverb setting. As the *audio film* cuts to Mary sitting in the living room, this shift of spaces is indicated by the removal of the reverb effect from the clock, the addition of the effect to the sounds coming from the kitchen and more explicitly by the processing applied to Jack's voice in the kitchen.

4.5. The Back Voice

The term *back voice* refers to those cases in which sound is processed to indicate that someone is speaking with his/her back to the listener [11]. The back voice was of significance in the design of the murder scene since, as Mary approaches Patrick with the leg of lamb he talks to her with his back against her so he does not see her approaching him with the murder weapon. Since there are no visual elements it was necessary to indicate this by applying reverberation to Patrick's voice and reducing the direct sound output.

4.6. Spatialization – Visualization and panning

In order to plan the process of spatialization the layouts of the living room, the kitchen and the bedroom were designed employing the 3D interior design software, *Interiors Professional*. See example of visualisation in Figure 1.
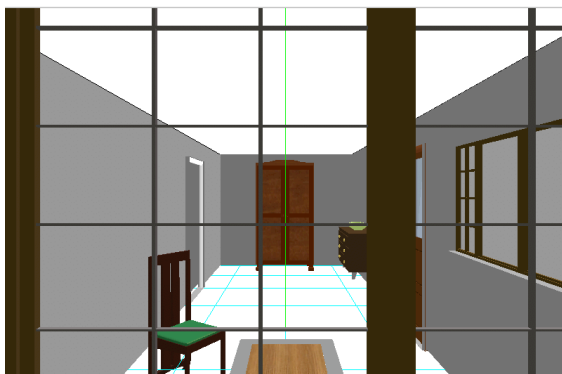

Figure 1. L*ayout of living room designed with Interiors Professional.*

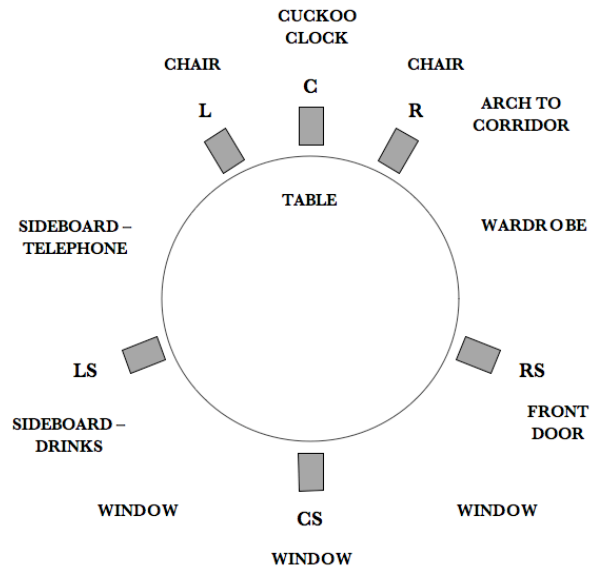Following the visualization of rooms the surround panning was planned. See Figure 2 for example.

Figure 2. *Layout of living room according to the 6.1 surround sound setup*

In the scenes in the street the outdoor sounds were spread through all the speakers to indicate that Mary can hear these sounds around her.

The most challenging aspect of the process of spatialization was to indicate the movement of the characters through different spaces. As the same channels are being used for every space and for the connections between them it was considered that listeners would have difficulties recognizing the transitions from one space to another.

In order to attempt to solve this problem each room entrance was assigned to a specific channel. These entrances were kept consistent throughout the *audio film*, and the spaces between them were considered as corridors connecting the rooms.
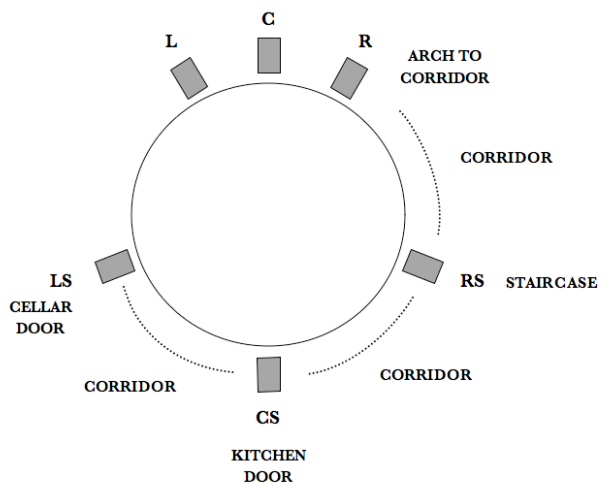

Figure 3. *Connections between spaces according to the 6.1 surround sound setup*

Another problem encountered was how to indicate that the characters were moving away from a specific space. In order to achieve this effect volume automation was employed, lowering the volume of specific soundmarks as the characters walk away from a room. However, the panning remains the same, failing to indicate the positions of the rooms in the house. For

example, when Mary exits the living room, the cuckoo clock is still panned to the C channel, even when she is in a corridor and the sound would have changed its position in relation to Mary.

In most cases the footsteps in the rooms were panned in a circular manner to make them more evident. This was done like this even though the visualization process might have suggested more direct paths for the characters' movements within the rooms.

Regarding Mary's footsteps going up and downstairs, the best option found was to use circular motion as if it was a spiral staircase.

When Mary is walking in the street since no fixed direction was determined circular motion was also employed.

In some cases footsteps were also sent to all the channels to indicate that the characters are approaching the centre of the room.

Music was sent to all the channels to create a sense of envelopment.

Mary's interior monologues were sent to all the channels to emphasize the fact that they are thoughts, so they do not need to be placed in an exact position.

5. AUDIENCE FEEDBACK

In order to test the effectiveness of the *audio film*, trial sessions were arranged in which 13 attendants were asked to listen to the work and complete a questionnaire that included questions regarding the perceived clarity of the *audio film*, plot understanding, the recognition of different characters, spaces and sound sources. Ideally, these sessions should have been attended by visually impaired people, however, due to the impossibility of getting volunteers, the testing stage was carried out with sighted people to get general feedback and analyze whether this format might be interesting for sighted people as well.

We first asked if the attendants had prior knowledge of the plot (see Figure 4) and the majority were not aware of the plot. Then, after listening to the *audio film*, we asked the attendants to summarise the plot.
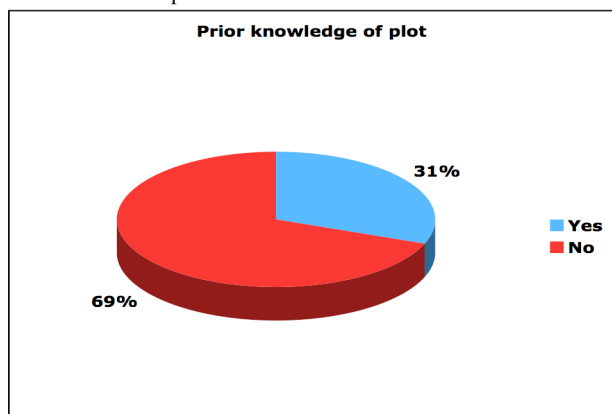

Figure 4. *Measure of prior knowledge of the plot*

Figure 5. *Accuracy of the summaries provided by the listeners who declared no prior knowledge of the plot*

Figure 5 shows the accuracy of the summaries provided by the listeners who declared no prior knowledge of the plot. This result is encouraging since the majority were able to understand the storyline. It must be mentioned that the volunteers that did not understand the plot were non-native speakers who mentioned having problems following the dialogues.

We measured the *clarity* of the audio film by asking the audience to name the characters they recognised, to names the rooms and spaces they recognised and to mention which sound elements helped in the identification.
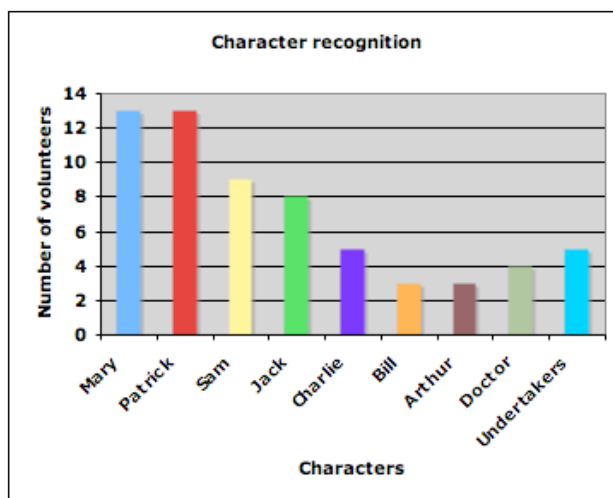

Figure 6. *Measure of perceived clarity of audio film*

The main characters (Mary and Patrick) were recognized by all the listeners (see Figure 6). The majority of the listeners recognised two other important characters Jack (the main policeman) and Sam (man at the grocery). However, the recognition of the remaining characters seems to have caused difficulties and none of the listeners were able to list all of them. This is probably due to the fact that as the plot moves forward the number of characters increases and it becomes more challenging to remember them distinctively.
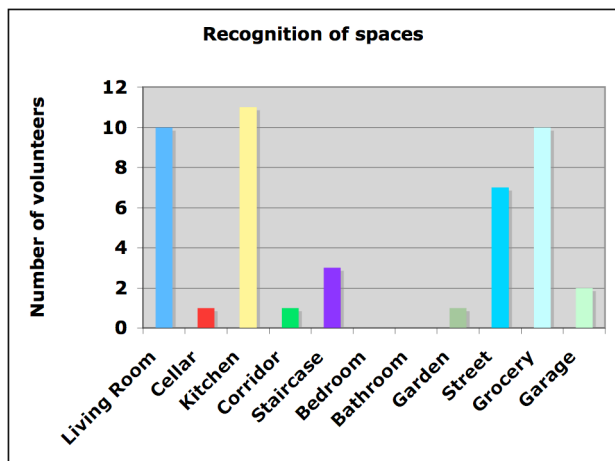
Figure 7. *Measure of recognition of spaces*

The spaces that were more widely recognized are those that have been assigned either soundmarks or sound effects that are typically related to those spaces.
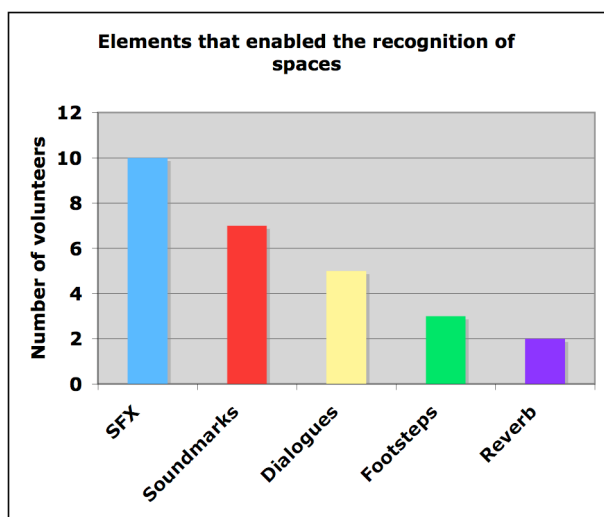

Figure 8. *Measure of elements to enabled the recognition of different spaces*

The two most prominent clues for space recognition mentioned in the questionnaires were the presence of characteristic sound effects such as cutlery in the kitchen, and the presence of soundmarks, such as the clock or street sounds.

6.   CONCLUSIONS

The design of *Lamb to the Slaughter* has demonstrated that it is possible to present a clear storyline solely through sound by employing sound effects, sound processing and surround sound to convey information eliminating the need of a narrator.

Sound effects were used both to represent actions and as soundmarks to help the listeners identify the different spaces included in the *audio film*. Also, artificial reverberation was employed to provide each space with a characteristic sound providing further differentiation.

As regards the use of surround sound, it could be noticed that it could be employed to suggest the layout of the spaces as well as indicate the movement of the characters.

Regarding the use of music, it can be used in the same way as films to indicate the characters' feelings, enhance tension and as a *leitmotif* to clarify the plot.

This project is only an initial insight on the potential of this format to convey a storyline without the need of visual elements and without the need of a narrator. Further work needs to be done to develop this format and analyze its viability. Research needs to be done on possible ways of finding sound equivalents of elements from the film medium such as different types of shots. More research also needs to be done on the portrayal of the connections between the different spaces, since although an attempt has been made to indicate these connections through imaginary corridors indicated through changes in the reverb settings, as well as by assigning the entrances to the spaces to different channels, this does not seem sufficient.

Moreover, a way of indicating differences in height should also be explored by the use of Ambisonics or a 10.2 surround sound setup. Additionally, it is necessary to explore the different challenges that this format would present when applied to different genres and aimed for different age groups.

Finally, it is of the utmost importance to test this format with visually impaired people to assess whether it could be used as a replacement of Audio Description and to explore the ways in which this format could be improved.

8.   REFERENCES

[1]   Richards, R. (1992) *A director's method for film and television*. London: Focal Press.

[2]   Piety, P. (2003), *Audio Description, A Visual Assistive Discourse: An Investigation into Language Used to Provide the Visually Disabled Access to Information in Electronic Texts*. Master's Thesis. Georgetown University.

[3]   Friberg, J. and Gärdenfors, D. (2004) 'Audio Games: new perspectives on game audio', *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, Singapore, 74: 148-154.

[4]   Ash, W. (1985) *The way to write radio drama*. London: Elm Tree Books.

[5]   Rattigan, D. (2002) *Theatre of Sound: Radio and the Dramatic Imagination*. Dublin: Carysfort Press.

[6]   Crisell, A. (1986) *Understanding Radio*. London: Methuen.

[7]   Levinson, J (1996), 'Film Music and Narrative Agency', pp.482-512 in Mast, G et al (2004) *Film Theory and Criticism: introductory readings*, 6th edition. New York: Oxford University Press.

[8]   Schreger, C. (1985) 'Altman, Dolby, and the Second Sound Revolution', pp. 348-355 in Weis, E. and Belton, J.

(1985) *Film sound: theory and practice*. New York: Columbia University Press.

[9] Sonnenschein, D. (2001) *Sound Design: the expressive power of music, voice and sound effects in cinema*. Studio City, Calif.: Michael Wiese Productions.

[10] Wishart, T. (1996) *On Sonic Art*. Amsterdam: Harwood Academic Publishers.

[11] Chion, M. (1994) *Audio-vision: sound on screen*. New York: Columbia University Press.

[12] Murphy, D. (2000) *Digital waveguide mesh topologies in room acoustics modelling*. PhD Thesis. The University of York.

[13] Blesser, B. and Salter, L. (2007) *Spaces speak, are you listening?* Cambridge: MIT