BA (Hons) Creative Music Production

Professional Project

By Sadhbh Bashford Rance

An Exploratory Study on How Binaural Sound Can Improve the Immersive Quality of Audio Storytelling

27th April 2022

Mr. Geoffrey Perrin

Dr. Brian Carty

Abstract: *The objective of this project is to explore the impact sound has on the immersive quality of audio storytelling. Research shows that firstly, music can induce an emotionally driven experience within a listener, secondly, that the use of foley, detailed soundscapes and sound design can be used to enhance narrations. Lastly, that binaural audio can evoke empathy within the listener. A project was undertaken wherein storytelling participants were recorded sharing stories. Which was followed by the recording of foley and soundscapes using a location recorder and the production of music and sound design was completed using both Protools and Ableton. All audio was then spatially placed within binaural environments using CSound. The results of this project show that the use of binaural audio brings forth a personal touch to the storytelling, which in turn heightens the listeners emotional engagement with the storyteller and the narration. It also shows that subtly-placed music and carefully thought-out foley can be effective in creating an immersive soundscape that will deepen the listener experience, especially when these soundscapes are binaural. Conclusions and observations made within this thesis highlight noteworthy findings regarding the immersive characteristics of foley, soundscapes and music, and how they can be utilised to bring another dimension to storytelling.*
Keywords: binaural, foley, music, soundscapes, storytelling

Table of Contents

Introduction

The topic of research is Immersive Binaural Storytelling. This research project explores the ways in which sound can be used to heighten the emotional response and immersive involvement of a listener when there are no visual queues. Participants will be recorded telling stories from their lives. They will be recorded on location recorders. Next the stories will be illustrated with foley noises and the use of soundscapes from both location recordings and additional sound design. One aim is to achieve the sensation that the listener is deep within the story that they are hearing by using binaural audio to place them spatially in the environment. This allows the listener to experience what the storyteller went through in those moments, making it personal to both the listener and the storyteller. This is an aspect of the project that is particularly powerful when looked at in relation to mortality and how these stories will live on long after the storyteller does not. It is also an important area of study because it is the future of storytelling, which is a significant part of many cultures. The use of the recreation of how the human ear spatially recognises sound sources is currently on the rise and can be seen in many aspects of the world around us and in turn makes what we listen to more real now than ever. For example, gaming, virtual reality, podcasts, and popular music all use binaural immersive technologies. Therefor this is a significant area of study as it merges the real world with the virtual world in a very humane way.

The goal of this project is to create and archive an immersive auditory storytelling experience which will be unique to each story. One that will allow the listener to get lost in a world that is illustrated by an array of sounds and evoked emotions. The sources that can be seen in the literature review were chosen to show examples of the different ways in which sound can be used in storytelling and to find evidence that music and carefully manipulated sounds can highlight certain feelings within the listener. Additionally, the production of a completely binaural soundscape allows for the storytelling experience to be elevated to another level of immersion, one as similar to the real experience as possible.

The personal style of storytelling, the significant use of binaural technologies, and the idea of immersion in this project are all linked by their tendency toward the personal. Firstly, binaural audio is used to emphasise empathy and emotion. How this is achieved is discussed in the literature review. Secondly, the type of storytelling used is personal as it captures peoples' unique ways of storytelling alongside their individual experiences. This makes the story immersive and personal for the listener as

it fully submerges them within the story. Additionally, the empathy-evoking binaural narration and the encapsulating binaural soundscapes emphasise the personal element further as the listener feels deeper involvement with the story and the environments that surround them**.**

Binaural recording is a way of recording audio that replicates the way in which we as humans hear. The procedure involves the placement of two microphones at ears distance to create a three-dimensional experience which will upon playback allow the listener to make a spatial placement of the audio source. Using coding software such as CSound allows for non-binaural recordings to be turned into binaural audio, spatially placing any audio source around the listener.

This interim thesis is comprised of a literature review and a methodology. The literature review surveyed a range of scholarly and authoritative sources in the fields of musicology, the use of binaural audio in storytelling, and the use of sounds to create an immersive auditory environment. The methodology for this project can be broken down into three distinct sections: pre-recording which is everything that had to be done before the recording of participants, the recording process itself where the story narrations are captured, and then lastly post-recording where elements such as soundscapes, music and foley will be added to increase the immersive quality of the project. The created auditory environments will be binaural in order to make them feel as real as possible.

Throughout this thesis the word ‘storyteller’ will be used to refer to the speaker who was recorded telling their story. The word ‘scene’ will be used to refer to sections within the stories.

Literature Review

This literature review is comprised of a mix of literature mediums spanning an array of topics that must be considered while creating this project, such as binaural audio, the power of music when conveying emotions, and the effectiveness of foley and sound design in creating soundscapes. Overall it will discuss the importance of audio, other than the narration, when it comes to telling a story. It will explore how foley, sound design, and music can be carefully balanced to create illustrative soundscapes that will enhance the listeners engagement to a story by making it feel more real, more personal, and overall more immersive. The literature highlights ways in which sound effects can be broken down into categories to help inform the creation of the soundscapes as to what sounds are important in telling the story. It will also look at different techniques and technologies available for creating binaural soundscapes.

The way in which humans hear is dependent on three factors. Firstly, phasing which is the emphasis, subtraction or cancellation of frequencies when audio is duplicated. Secondly, the direction in which the sound is heard, and thirdly, the physical separation of the two ears, being the head. Binaural audio replicates the way our ears work. As discussed in the journal article by Shane Hoose, "Creating Immersive Listening Experiences with Binaural Recording Techniques",[1] there are three main recording techniques that can be used to achieve this all which try to mimic the human anatomy of the ear when it comes to head-related transfer functions. Head-related transfer functions distinguish how an ear experiences a sound source within a three-dimensional space.

Firstly, the dummy head technique relies upon the construction of a fake human head with omnidirectional microphones, microphones that pick up sound equally from all directions, within the ears of the dummy head. When creating this dummy head the measurements are extremely important in order to get it to work binaurally. It is usually the average-sized head, nose, ear canal and shoulder sizes used so that they work best at replicating the majority of peoples' hearing.

Secondly, the baffled technique also uses omnidirectional microphones, but instead of being mounted in the ear of a dummy they are placed at ear distance apart and separated by an acoustic baffle made from a sound absorbent material. Although this replicates the physical separation of the ears by the head it does not mimic the shape of the human ear and how they can effect the way in

[1] Hoose, Shane. "Creating Immersive Listening Experiences with Binaural Recording Techniques." *College Music Symposium*, vol. 55, College Music Society, 2015

which we perceive sound. Recording the audio this way does not give such an accurate representation of human hearing but one upside of choosing the baffled technique is that upon playback of the binaural audio over loudspeakers it can sound rather spacious, something that does not transcribe well with the other two techniques.

Thirdly, head worn microphones or in-ear microphones are microphones that are placed within the human ear and replicate hearing accordingly. Recording the audio this way allows for head movements which can be very helpful for the listener in spatially placing sounds in the environment surrounding them. This freedom of movement is one of the reasons why binaural in-ear microphones will be considered when recording the stories within the project.

If sounds are recorded using a microphone setup whose positioning mimics that of the human ears, then upon playback through headphones the three-dimensional placement of the sound will be preserved and allow for the listener to be within the recorded environment. This is also possible to achieve with the use of an audio programming software called CSound. In CSound binaural opcodes allow for mono sound sources to be taken and placed at any angle around the listener, creating a three-dimensional, virtual environment. This means the audio does not have to be recorded using binaural techniques and instead can be recorded using a stereo or mono setup. CSound opcodes 'hrtfmove' and 'hrtfstat', which are techniques explained by Brian Carty in the article "hrtfmove, hrtfstat, hrtfmove2: Using the New HRTF Opcodes", help to replicate the space the story occurred in by adding audio sources that can either stay static (hrtfstat) at an x and y angle around the listener, or can move (hrtfmove) at angles around the listeners head.[2]

A 2019 study "Music Emotion and Visual Imagery: Where Are We Now?" By Küssner, Erola and Fujioka, on the connection between music, emotion and visual imagery found that music has the power to induce an emotionally-driven visual experience in the listener.[3] Participants in this study reported seeing mental images of landscapes, people dancing and other obscure imagery such as shapes and colours while listening to a piece of music. This research is important because it highlights music's ability evoke a personal internal experience. For this reason, music will be incorporated into this project to enhance the storytelling experience for the listener by heightening their emotional engagement.

---

[2] Carty, Brian. "hrtfmove, hrtfstat, hrtfmove2: Using the New HRTF Opcodes" *CSOUND JOURNAL*: Issue 9. 2008

[3] Küssner, M. B., Erola, T., Fujioka, T. "Music, Emotion, and Visual Imagery: Where Are We Now?" *Psychomusicology: Music, Mind, and Brain,* vol. 29, no. 2-3, 2019, pp. 59-61.

Interestingly, the research also highlighted that the visual response to the piece of music was unique to each individual. The images that were visualised spanned a huge range, and would most often relate to personal experiences. Therefore the use of subtly placed music along with binaural recording, foley sounds and soundscapes in this project will enhance the listener's emotional experience of the narratives while the vocal storytelling itself will remain the sole focus. Additionally the ability of the music to recall visions of episodic memories in the listener as pointed out by Küssner, et al in the article, will help listeners of this project to relate to the stories they are hearing.

Pauletto and Lopez compiled a study in 2010 titled "The sound machine: a study in storytelling through sound design" which surrounds the idea that narration is not a necessary factor in storytelling and that through sound alone they would be able to tell a comprehensive story.[4] Their findings were that this could be achieved through a creative use of sound design. An important part of their research is their breakdown of the audio into three sections: Ambience tracks (birds singing or the screams of children in a playground), Sound effects (footsteps, creaky floorboards and phone ringtones), and Internal sounds (such as breathing). Each is useful for this project when it comes to creating detailed soundscapes to bring the stories alive. Their aim was to be able to produce a story that would be narrated with no dialogue, only sounds. Whereas with this project, the words of the storyteller will be the main focus. So although the final product of their research is completely opposite to this project in that they eliminate narrative and this project will centre around it, their study showed the true power of sound and how it can be used to bring a whole other dimension to the world of the storytelling.

In 1978 BBC 3 aired an innovative new radio play by Andrew Sachs, "The Revenge", which was experimental at the time.[5] It was a thriller that was told completely through sounds alone, without the help of any coherent speech. This radio play is important to highlight the ability and extent of control sound has to actively effect what a listener is perceiving, using sounds to illustrate, which is something very important when bringing the stories in this project to life. The radio play achieved what was intended, it tells the story of a man escaping from prison following him through fields, and bodies of water, evoking tension as you picture what could be waiting for him next. The Revenge was recorded on location using the binaural stereo techniques available at the time. This use of location

---

4 Pauletto, S., Lopez, M. "The sound machine: a study in storytelling through sound design." *AM '10: Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, no.9, 2010, pp. 1-8.*

5 Sachs, Andrew. *The Revenge. YouTube,* uploaded by RandomRadioJottings, 4 Dec. 2016

recorded sounds is something that is intended to be called on frequently in this project, taking advantage of the environments already around us.

Darkest Night is a 2018 storytelling podcast, by L.Pace, where in each episode the listener is placed inside different scenarios.[6] The goal was to achieve an immersive experience, brought to life with the use of binaural audio, foley and different soundscapes and sound designs. Being a horror story podcast the use of binaural audio in this context is incredibly effective at enhancing the scare quality and thus the experience for the listener. On listening to the binaural recording the listener is placed within the story which evokes empathy for the characters and storyline as they can see themselves in their shoes, something thought to be very important when deciding to use binaural recording. If you take one episode for an example, Goat Trees, there is an almost constant use of ambient tracks and sound designs to create a subtle yet effective atmosphere which heightens the evoked emotions. The use of foley and sound effects, such as footsteps, breaking sticks and goats bleating, illustrates what is happening in the environment around the listener. When referring to this project, there are multiple ways in which it will differ from Darkest Night, mostly regarding the way in which the monologue will be recorded. Recording using binaural in-ear microphones or binaural techniques in Csound will allow for the narration to be moved around the environment and in turn making it more immersive for the listeners of the project.

Award winning podcast, "*Have You Heard George's Podcast?"*, from *George the Poet* also uses binaural audio, sound design, and soundscapes to enhance the listeners experience with the podcast narrative.[7] Although both podcasts use sound to emphasise their content, they differ in how they use sound to evoke emotions. *Darkest Nights* focuses on creating tension and adding to the fear factor while *Have You Heard George's Podcast?* uses binaural and other immersive audio techniques to experiment with the impact of sound. The narratives being told in George the Poet's podcast are very intimate and they use the audio to drive emotions home to the listener, painting a mental picture for them with the sounds and what each one represents. But just like *Darkest Nights*, *Have You Heard George's Podcast?* is going to differ from this project in how immersive sound is being used to heighten emotions within narratives. In both these examples of podcasts the soundscapes can be at times quite sparse and reliant on the narrative to drive the imagery, whereas with this project the narrative will of course be used to drive the imagery and influence it but ideally the story would not

---

[6] Pace, L. "Goat Trees." *Darkest Night from The Paragon Collective,* 31 Oct. 2018

[7] Poet, George the. "Vibrations." *Have You Heard George's Podcast?,* 2 Sept. 2021

rely on the narration to keep the pace, and instead with the subtraction of the narration for a period of time the detailed and ever changing soundscapes would still succeed in painting some sort of storyline in themselves, adding to the overall immersive power of the finished product.

Twenty Thousand Hertz is a 2020 podcast that takes a look at the history of some of the most iconic and interesting sounds that we hear today, hosted by D. Taylor.[8] Each episode is home to a handful of examples of how sound design can be effectively integrated with the content of a purely audio based interaction. Its aim is to enhance the experience and make it more stimulating. The execution of the sound designs is effective in drawing the listeners in. This podcast is educational in two ways: Firstly, it is a good example of the effective use of sound design in storytelling, and secondly the content of the podcast itself discusses topics such as musicology and how music can change moods and poses questions such as "is it us or the music that is in control" when it comes to emotions. This can be heard in the episode titled "Good Vibes".

In conclusion, there are three areas of interest to take away from the research. Firstly, the importance of music to heighten the listeners emotional engagement with the stories. Secondly, the importance in the use of ambient tracks, sounds effects and internal sounds to recreate an environment for the listener, and how it can be strategically used alongside the narratives to completely immerse the listener in the world. Finally, the binaural audio of the storyteller is important when looking to evoke empathy from the listener because mentally it spatially places them there in the story. Therefore the use of subtly placed music with binaural recording, foley sounds, and soundscapes in this project will be essential in enhancing the listeners overall emotional experience of the narratives.

8 Taylor, D. "Good Vibes." Twenty Thousand Hertz, 5 Feb. 2020

Methodology

As discussed previously the aim of this project is to create an immersive three-dimensional auditory storytelling experience, one that will employ certain techniques that will heighten the listener's emotional engagement with the stories recorded. These techniques and their planned uses are outlined below.

This project follows a methodology that can be divided into three key stages. Pre-recording, recording and post-recording. When looking at the pre-recording stage, it can be broken down into the further sections of 'equipment' and 'participants'. When picking a microphone for this project it must be kept in mind that binaural audio plays an important role in the production of immersive experiences. For the recordings, the consideration of binaural in-ear microphones was taken into account. Although after careful consideration of its effectiveness alongside the binaural soundscapes that would be created in CSound, it was feared that their acoustic characteristic differences may clash and result in the overall experience being less immersive, and so using binaural in ear mics for recording the stories was decided against. Instead, a Zoom H5 location recorder was chosen as the primary recording device for both narrations and all other recorded sounds. The makes and models of the equipment are irrelevant and the project could be successfully completed with other makes, however they are named for reference purposes. Tripods and headphones were also attained to aid in the recording process. Before anything could be undergone regarding the production of the project, the enlisting of participants happened. Social media and word of mouth was used to help find a group of people from a variety of ages and backgrounds who would like their stories heard.

When it comes to the recording process, the location of the recording took place wherever the storytelling participant felt most at ease, be it in their home, or somewhere else. The chosen location was permitted under the current Covid-19 guidelines at the time and was preferably somewhere with little background noise. Although it was preferred to record in a location with minimum background noise, the recording should not sound too dead either, so subtle background noise and room ambience was acceptable. On arrival to the recording location, it was important to make sure the participant had everything that they needed, water for when the recording was underway, somewhere where they could comfortably tell the story, and an understanding that there is no pressure and that they could take their time and perform at their own pace as they feel comfortable. Before and during the recording the microphone placements were taken into careful consideration. When placing the microphone,

proximity effect was carefully considered. The Zoom 5 location recorder used its stereo x/y microphone placements and was placed between six to twelve inches away from the storyteller and mounted on a tripod. The use of in-ear microphones was considered as this would have allowed the person wearing them to move around whenever they felt necessary. This is something that would directionally place the listener in the space that the story is being told. If the sound recordist decided to move around while recording, proximity effect should be carefully considered. However, in-ear microphones were not used for recording as this effect could be replicated post recording and was therefore deemed unnecessary.

The aim of this project is to immerse the listener in the story being narrated, so the more detail from the storyteller the better. If the participants were not giving enough detail in their stories, they were asked questions to prompt more detail and description. For example, the participants were asked things like, “how did that make you feel?” Following the recording, some storytellers were asked for details about the environments in which their story occurred. This could be the regarding the shape of a room, whether there were carpets on the floor, where the doors were, or where any important centrepiece of their story was in relation to them. This is in order to form the best understanding of how the surroundings might look and allowed for a more accurate representation of their story.

Once the recording was complete, post recording began. The first step post recording was cutting and cleaning up the audio. Editing out any prompts or unwanted noises. Listening through the audio carefully, picking out anywhere the use of soundscapes could work, highlighting provocative key phrases. The story was then played to someone who has yet to hear it, with the idea that a new pair of ears would help to pick up anything important that may have been missed. Most of the stories required a certain amount of research before the creation of their soundscapes could begin. This could include finding out what sounds were around at the time of the story, and is important in creating an accurate depiction. An example of this would be the emergency services sirens that were around in 1950’s England. After doing research it was discovered that they had a certain type of bell attached to the vehicles instead of a siren. The soundscapes were recorded using the location recorder. The addition of any music to help illustrate evoked emotions then took place and was added into the stories in any natural pauses or where it complimented the narrative. Pauses were lengthened where appropriate to make space for illustrative audio. The music was created using physical instruments, recorded using Protools, and also with virtual instruments and sound design using Ableton Live.

Before creating any of the music, research was done into how emotion is projected through the human voice, which was helpful in the creation of music that would hopefully project these desired emotions.

Lastly, there was the addition of any foley that was felt necessary. The foley was mostly recorded using the location recorder but the use of pre existing sound effects was considered when any of the sounds needed were too difficult to record. As the library of sounds for each story grew, they were added into a virtual three-dimensional environment using the coding software CSound, which was achieved with the use of opcodes such as hrtfmove and hrtfstat. These opcodes created a binaural environment by spatially placing sounds at angles around the listeners head, which illustrated and enhanced the story as the soundscapes and music weaved through the narrative and the virtual environment. Extra opcodes such as granule and hrtfreverb were utilised to produce more complex and textural elements that were then used in the soundscapes. It must also be noted that when creating the binaural audio an ear size needed to be chosen. For this project the average ear size was used.

The final step of the methodology for each individual story was the allocation of a title. The titles were chosen by listening through the narration and picking sentences or phrases that stood out for either being particularly visually strong, or which showed the character of the storyteller.

The rationale for the above methodology will now be discussed. The reasoning behind choosing to record the stories using a stereo set up and later making them binaural using Csound, instead of initially recording them binaurally with in ear mics, was because when recording the story narration with in-ear binaural mics this would include the acoustic characteristics of the space that the story was initially recorded in. There was a worry that these acoustics would not compliment the environment in which the story is meant to be set and would risk pulling the listener out of the desired environment and further diminishing the immersive quality. Instead, by using the stereo recording, and later binaurally placing it within the stories binaural environment using Csound opcodes, this would avoid any acoustic clashes while still having the empathy evoking characteristics that binaural audio holds. By using the stereo recording this also left room for more creative freedom for future decisions regarding the placement and movement that the story narration may take within the three-dimensional environment of the finished piece. Upon hearing the immersive binaural storytelling, the listener will be placed within the story, and this will help to evoke empathy for the storyteller and the storyline.

Binaural audio replicates how people hear, which for this project will allow the listener to directionally place themselves alongside the storyteller within the soundscape of their experience. Podcasts such as *Darkest Night* (Pace, 2018) use this method of binaural storytelling quite effectively and serve as inspiration for this project. The use of a location recorder was chosen for its diversity of uses and for recording a variety of elements of the project: to record the story narrations, soundscapes and foley discretely and with ease, as well as recording any instruments to be used in the production of the music. The tripod was used to assist the location recorder and provide stability.

Social media was partially used to enlist participants because it is an affective way of bringing the project to people's attention. However, the enlisting of participants was mainly done by word of mouth, which is fitting to the nature of the project. A variety of ages and backgrounds were selected to provide a wider and more diverse range of stories, from different times in the past decades, which were told by a variety of different kinds of storytellers, all telling their story in their own unique way, free from the restraints of any preconceptions the listener might have of them as they will be unaware of anything about the storyteller except what they choose to share.

Recording the storyteller in a space that that they feel most at ease was decided upon because it would be the setting most likely to result in the best story: one where the teller feels most confident and delivers the best performance possible. A space with little background noise as to avoid any unwanted noises pulling the listener out of the experience, but also a space that is not too dead so to keep some character to the audio. The use of location recorders and other lightweight handheld equipment meant that this style of recording in various places was very accessible and did not require much time to set up. The positioning of the stereo microphones when recording the storyteller was kept consistent between each story. Any wanted movements around the listener was added at a later stage when CSound was being used to create the binaural environments. The microphones were placed six to twelve inches away from the storyteller. This was for a reason similar to proximity effect. As a person's voice gets closer to a microphone it can boost low frequencies and in-turn become overly bassy. This was important to keep in mind particularly when recording male storytellers**.** Microphone placements were important when capturing a stereo recording. For this project x/y microphone placement were chosen because it is ideal for close microphone setups. It also gives a stereo image that is not very wide. This is preferred when trying to avoid surrounding noises. Other microphone positions are equally suitable but the x/y microphones come as part of the ZoomH5, solidifying the decision to use that specific microphone placement.

Getting the balance between the placements of the listener and the storyteller is important as it will place the listener in the scenario of the story while also feeling like the storyteller is there with them. As heard in the radio play 'The Revenge' (Sachs, 2016) stories can be effectively told entirely through sound and without any narrative. The aim of asking the storyteller questions to prompt detail and description was to assist the production of the soundscapes and foley that emphasised the surroundings while simultaneously trying to capture a feeling as similar to that of the original experience as possible.

During the post recording process, cleaning up the audio was done to achieve a smoother recording, cutting out any prompts or unwanted noises that would draw the listener out of the world and risk making the storytelling experience less immersive. This also included lengthening any pauses in places where music or soundscapes benefited from being focused on. The clean audio was listened through, picking out any provocative phrases that evoked a feeling of the surroundings. This was where there was a need for a soundscape, to paint the auditory picture of the story's environments. An outsiders perspective was taken into consideration to listen through to see if there were any key phrases that were overlooked and would be suitable for the soundscape. As the finished project is for anyone to listen to, and each person perceives things in different ways, it was beneficial to get the input and feedback from an unbiased listener during this phase of production. The importance of researching certain aspects of the stories is particularly important when there are elements to the story that are unfamiliar, or from a time period prior to the present day. By gathering information about technologies at the time, or sounds associated with specific locations it makes for a more accurate/ believable and immersive storytelling experience.

The addition of music was used to highlight any evoked emotions. They were placed in the stories natural pauses to help set the scene, or placed behind the narration in order to emphasis the emotions and motions of what is being said, working with the storyteller without taking focus away from them. Research was undertaken to explore how emotions are projected through the human voice. This was in the hope that what was observed could be translated into the way that the music is created. Ericson (2022) in *A Voice of Frustration*, looks at volume and how its fluctuation can have power over the emotions that we perceive in speech. During a casual conversation peoples voices tend to remain at a certain level with only slight variations. When angered the voice tends to get lower and louder, while when happiness is present in the voice it tends to increase in both pitch and volume, and when

someone is frustrated or the frustration is wanted to be portrayed in music then there is a want for roughness and intense sounds that lack fluidity by decreased pitch and increased volume, using what was referred to as a ‘chest voice’.

When creating any music, Protools was used for tracking as it is a preferred audio workstation for recording instruments and mixing. This was also used for the editing of the initial narrations. Ableton Live was used for some of the musical compositions as well as sound designs, due to its suitability for soundscape design. Most of the foley was recorded using the location recorder as it is practical to transport, versatile for using in different types of locations and lastly because this project aimed to use as much original sound as possible. The use of any pre-existing sound effects and foley noises was considered for use if it was not possible to record those sounds first hand and on location. It was kept in mind that the use of too much foley could potentially take away the sense of realism that this project aimed to achieve and as such it was used sparingly.

CSound was the coding software chosen for use in this project as it is a free computer programming language for audio programming and it houses binaural opcodes that were useful in the binaural reconstruction of the narratives. The binaural opcodes that were used throughout this project to create the encapsulating environments were: ‘hrtfstat’, which allowed for an audio source to be placed statically at any x and y angles around the listeners head, ‘hrtfmove’, which facilitates the movement of an audio source around the listeners head, and ‘hrtfreverb’ which is an audio effect that adds binaural reverb that can be personalised to each sound. As well as binaural functions it also allows for other audio effects such as granule, which is a granular synthesis opcode that was used to create thick textural sound designs. Granular synthesis is when a piece of audio is broken up into small fragments called grains and then these grains are rearranged and played back, forming a cloud. These grains can be layered on top of each other, and played at different speeds, volumes and frequencies. Granular synthesis can be a very effective sound design for creating soundscapes. Having completely binaural soundscapes to go alongside the story narrations allows for deeper immersion and a more real sensation when listening back and one that will draw the listener right into the emotional core of what is being told.

It was in this stage of the methodology where the volume levels of each auditory element were decided. The first level to be inputted was that of the narration. Once the narration level had been picked for the first story, each narration volume for the stories to follow used the first narrations

volume as a reference. This was in order to maintain a volume that would stay as close as possible throughout each finished piece. Once the volume level for the narration was set, then every other added sound's volume would be based off the narration, so that they worked together. When creating the binaural audio an ear size must be chosen, this is so CSound knows how to manipulate sounds in order to make them sound directional and placed around the listeners head appropriately. Everyones ears are different, and their individual shapes and sizes impact on how they hear. For this reason an average ear size was chosen for this project in order to accommodate accuracy for the most amount of people possible.

When choosing a title for each story, evocative phrases that emphasised strong visuals or the character of the storyteller were ideal as they both hold significance/links with the goals of this project. Phrases that hold strong visuals would work well as that is a main way in which illustrative audio was decided upon. For example, one of the titles was “A cloud of missiles”. This was chosen as it was a phrase that was not only mentioned a few times within the story but also held an importance within the story, not only to its narrative but also to the foley and soundscapes that were used to paint the auditory picture. The other possibility for titles was to pick phrases that showed the character of the storyteller. By picking titles in this manner it emphasised the personal element to the story, and the goal of capturing the storytellers unique way of storytelling.

The majority of the chosen method has been trial and error on a previous preparation project with many similarities throughout all three stages of the project: pre-recording, recording and post-recording. Due to this previous experience, an accurate estimation of the immersive quality that is possible when following this design was achieved, that of a high standard and one which successfully achieved the desired emotional effect. The main difference with the current project is the use of completely binaural soundscapes. The addition of this encapsulating audio will only further boost the immersive experience and heighten the listeners emotional involvement with what they are hearing.

Potential limitations should be taken into consideration. A limitation very current to the times in which this project went underway is the Covid-19 pandemic. This is where the most pressing downfall of the project lies: limitation on the variety of participants available due to societal restrictions. During the recording stage of the project the pandemic restrictions fluctuated rapidly and without much warning, meaning that some participants were hard to meet with in person. This resulted in a smaller variety in ages of storytellers as it is the older generation who became more cautious and

reluctant to meet in person. Health Care guidelines were followed strictly. The second limitation regarding participants is what they were willing to share. However, when a storyteller's narration was lacking in description, despite the prompt given, then this was treated not as a limitation but as a chance to channel more avenues of creativity in terms of finding a way to make the sparse narration as real and as immersive as possible while still remaining true to the story being told.

Analysis

Observations made about this project can be broken down into four key sections: the story recording, the use of foley and soundscapes to highlight environments, and music which is subtly placed to emphasise certain images and emotions connected to the story. These three sections hold the foundations for the fourth section; the immersive quality of the project as a whole, and the execution of the binaural audio which was significant in emphasising the personal element to these stories.

Firstly, when reflecting upon the storytelling narratives and their recordings it must be considered that despite the fact that overall they elicit what was intended of them in regards to capturing peoples personal experiences, that does not mean that they were captured without issues. At times they were recorded in rooms that were not ideal and needed editing to subtract out any resonant frequencies. In rare cases the storyteller wanted to do several takes of the same story so this involved editing takes together to create the best version possible, although this was kept to a minimum as to not take away from the natural flow of the storytelling. It was also at times challenging to portray what was being said. Particularly in one of the stories where the storyteller had impaired speech that could be difficult to understand. Because of this it was heavily up to the additional audio such as foley and music to guide the listener through the narration, something that the literature review proved possible. Despite these difficulties in understanding, this is exactly what was wanted from the project, to capture the unique and personal ways of storytelling. Recording the stories using a stereo set up was successful in providing freedom when choosing the best placement of the narration within the stories environment, it made way for creative experimentation when creating the flow of the story and its soundscapes. The overall variety of stories that were told was strong, providing emotional diversity. This was also interesting when seeing what people decided to share. Although having more participants would give for an even greater variety of emotions not just in the stories but also the ways in which they are told.

Secondly, when looking at the use of foley sounds and soundscapes to highlight environments, throughout this project it can be seen that these were executed effectively and with precision. Close attention was paid to the placement of these sounds in relation to the listener as they navigate their way through the narrative. Potentially what made the soundscapes so visual was the inclusion of sounds that were not necessarily mentioned by the storyteller in their story, but were instead inserted as a means to help set the scene and enable the listener to envision the atmosphere surrounding the moment. This was an element that was particularly powerful in the case mentioned in the previous

paragraph where it talks about the storyteller who was at times hard to understand. During these moments foley was key to enforcing the scene and its narrative. Overall in the stories the foley and soundscapes are used almost constantly but also sparingly, meaning that while there is a steady presence of soundscapes throughout, they do not attempt to illustrate every word. Overuse of foley would have potentially taken away from the aspect of realism and resulted in a less immersive listener experience.

Thirdly, the successful use of subtly placed music to emphasise images and emotions within the story is slightly more difficult to measure. The effectiveness tends to fluctuate depending on the scene within a story. There are few areas where the music and sound design used is questionable due to the subjective nature of the emotions in the narrative. For example, calm background music used in a story which narratively is objectively not calm, may seem like an uninformed choice but was used to create immersion in a situation where emotions may be conflicting. Certain emotions were emphasised during the music and sound design that may not be reflective of each and every listener experience as each listener brings their own subjective meaning to a piece. However, these instances are occasional, and more often than not the music and sound design successfully goes hand-in-hand with the emotionality of the narration, helping to vividly paint a picture for the listener. This was usually in scenes where there was a distinct emotion or feeling within the narrative. Arguably at times the volume level of the music is not ideal as it takes emphasis away from other audio elements, and although it does not pose that big of an issue, is a change that could be considered in certain scenes.

Lastly, and most importantly, the immersive quality of the project as a whole must be analysed, taking into consideration the three sections outlined above but also with great focus on the binaural aspect to the storytelling. In this project the most important aspect is the immersive quality that it holds. The overall aim of this thesis project was to explore how sounds, particularly when binaural, can be used to improve the immersive quality of audio storytelling. Through the use of binaural audio, subtly-placed music, foley, and soundscapes, this project was successful in increasing the immersive quality of the audio stories and thus the listener experience. The use of CSound was both a practical and efficient way of putting together individual sounds and creating captivating binaural soundscapes.

In conclusion of this analysis, it is worth noting that the concept of immersion usually conjures up images of largeness and engulfment. However, this project is a good example of the power of smaller and more subtle uses of immersive techniques in sound design. This subtleness, combined with nuance in design, can be the determining factor in just how successful the immersive element really is.

Discussion

The previous analyses showed that while the thesis project has been successful in exploring the immersive qualities of sound design in audio storytelling, there remains room for improvement and learning. This project design was appropriate for answering the research question as it used auditory elements such as foley and sound design to non-visually paint a picture while music and the binaural aspect helped to emphasise and evoke emotions.

To begin with, while undertaking the auditory illustration of these stories it was important to envision the intention of the storyteller and perhaps to understand what they were trying to share with the listener through their choice of words. For example, was the storyteller trying to impart humour? Or to frighten the listener? Or to comfort them? Because the storyteller was not questioned on the intent of their story, the stories were listened back to objectively and sounds were then designed that reflected this objectivity. This process was unique to each story and involved the breakdown of the narrative intent so that it could be utilised in the decision making and selection of the chosen sounds, especially the music and sound design. This highlights the importance of objectivity when extracting narrative intent as a necessary skill.

Following on from discussion in the analysis, there are two things that are important to keep in mind when using sounds to add to the immersive quality of a narrative. Firstly, that subtleness is incredibly powerful and effective, and it can be the leading factor in determining whether or not the addition of sounds is adding to or taking away from the goal of immersion. The correct balance can be hard to attain and can take a lot of trial and error to get the levels right. For this project a lot of care was taken to keep the emphasis focused upon the storytelling without the presence of the sounds being too prominent and drawing attention away from the story itself. One of the project's aims was to effectively illustrate through auditory means what the storyteller was saying, and to avoid telling the story for them unless necessary. This relates back to the skill of objectivity in narrative extraction mentioned above. Secondly, that volume is an important subset of subtlety and should be given particular consideration in regards to the musical components. It was important to find the correct balance between overbearing and unnoticeable music. This also went for foley sounds and the general soundscapes. It was important to remain focused on the job at hand which was to enhance the storytelling and not to drown it out or replace it with sound. Listening skills were required to listen out for these balances while reviewing the work.

Again, the aim of this project was to explore how sound can be used to improve the immersive quality of audio storytelling; gathering stories from people of a variety of ages and backgrounds to get a more diverse array of narratives. Unfortunately due to the coronavirus pandemic restrictions in place at the start time of production, and the underestimation of how long each story would take to illustrate, the scale of the project was unable to reach its maximum potential and had to be reduced in size. This meant that the project could not document the full array of narratives that was originally planned and subsequently was not as culturally rich as hoped.

Another element of the project that fell short between the planning stages and the execution stages was the uses of foley sounds that were recorded completely firsthand. The original intention was to record all foley sounds firsthand using a location recorder, and for the most part this was achieved, however, during the post-recording process some of the needed sounds were hard to attain or replicate with what was available. In the place of certain recordings, online sources such as Soundsnap, a copyright free website with a library of sound effects, were used to carry out certain foley noises that were unable to be recorded first hand.

Despite these set backs the project was able to achieve its main objective of using binaural audio and the addition of carefully selected and placed sounds to improve the immersive quality of the storytelling, as outlined in the analysis.

Conclusion

To conclude, the research undertaken as part of this thesis showed that music can induce an emotionally-driven experience within a listener, that sound design enhances narration through the use of sound effects, foley and detailed soundscapes, and finally that binaural audio allow listeners to empathise with the storyteller and to put themselves in the storyteller experiences. The aim of the practical element of this project was to put this learning into practice while recording stories told firsthand from selected participants.

Participant's stories were recorded using a location recorder which was later used to record foley sounds and sounds for the soundscapes. Music that complimented the narratives was recorded and produced using both physical instruments and digital instruments via programs such as Protools and Ableton. The speech recordings, foley, soundscapes and music were compiled together in CSound, creating completely binaural immersive audio storytelling experiences using binaural opcodes available in CSound.

Overall, the project was successful in its goal to use sound as a means to enhance the immersive quality of storytelling. However, there are some improvements of note; firstly that the binaural aspect is more effective when movement and diversity of placement is used as opposed to solely static placements. Secondly, the use of certain pieces of music and sound design were not as objective as hoped. However, most of the music and sound design was effective in subtly emphasising the intended emotions at the correct times.

Where the project excelled in creating an immersive experience was through the use of foley and soundscapes. The balance of foley and the soundscapes used was able to add realism to the narratives. The most important piece of learning taken from this project is the notion that subtlety is effective, and the overuse of any tools such as music or foley can be detrimental to the finished piece by contradicting the immersive experience and taking away from the aspect of realism. It is important to remain focused on the task at hand which is to enhance the narrative without stealing from it or drowning it out, remembering that it is the storytellers story.

It must be noted that this immersive binaural storytelling project would have benefited immensely from completely taking place in a non-pandemic world, allowing for access to a vaster display of stories. It is a project with no definite ending and with countless amounts of possibilities for continuation, one that will be developed on into the future, taking into account what was learned

during this initial project. With the overall main goal being to get a bigger number of participants, with a wider range of diversity in storyteller but also emotions. In an ideal world these stories would also be put on a platform where they could benefit from having a bigger audience, and in turn more participants.

To conclude, the value of this area of study must be highlighted. The creative use of sound offers a chance for today's technological advances to work alongside the deep-rooted historical art of storytelling; something that holds a significant importance in many cultures around the world. What this study explored was how binaural sound techniques can merge the art of storytelling with modern technologies by means of its immersive quality.

Bibliography

Carty, Brian. “hrtfmove, hrtfstat, hrftmove2: Using the New HRTF Opcodes” *CSOUND JOURNAL: Issue 9. 2008,* http://www.csoundjournal.com/issue9/newHRTFOpcodes.html

Ericson, Ashley. *A Voice of Frustration.* Sonic Dictionary. https://sonicdictionary.duke.edu/exhibits-pages/voice-frustration-ashley-ericson [Accessed 23rd Jan 2022]

Hoose, Shane. “Creating Immersive Listening Experiences with Binaural Recording Techniques.” *College Music Symposium*, vol. 55, College Music Society, 2015, https://symposium.music.org/index.php/55/item/10863-creating-immersive-listening-experiences-with-binaural-recording-techniques

Küssner, M. B., Erola, T., Fujioka, T. “Music, Emotion, and Visual Imagery: Where Are We Now?” *Psychomusicology: Music, Mind, and Brain,* vol. 29, no. 2-3, 2019, pp. 59-61. http://dx.doi.org/10.1037/pmu0000245

Pace, L. “Goat Trees.” *Darkest Night from The Paragon Collective,* 31 Oct. 2018, http://www.darkestnightpod.com/season-3

Pauletto, S., Lopez, M. “The sound machine: a study in storytelling through sound design.” *AM ’10: Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, no.9, 2010, pp. 1-8.* https://dl.acm.org/doi/abs/10.1145/1859799.1859808

Poet, George the. (2021). “Vibrations.” *Have You Heard George’s Podcast?,* 2 Sept. 2021, https://www.bbc.co.uk/programmes/p07915kd/episodes/downloads

Sachs, Andrew. *The Revenge. YouTube,* uploaded by RandomRadioJottings, 4 Dec. 2016, https://www.youtube.com/watch?v=tsAdYGdZc88

Taylor, D. “Good Vibes.” Twenty Thousand Hertz, 5 Feb. 2020, https://www.20k.org/episodes/goodvibes