Dalarna University College
Sound and Music Production
180 ECTS Credits Bachelor Programme
Bachelor Degree Thesis

Game Audio in Audio Games

Towards a Theory on the Roles and Functions of Sound in Audio Games

---

Written by Rickard Åsén

Supervisor: Johnny Wingstedt
Course coordinator: Toivo Burlin
Examiner: Jacob Derkert
Publishing year: 2013

Abstract

For the past few decades, researchers have increased our understanding of how sound functions within various audio–visual media formats. With a different focus in mind, this study aims to identify the roles and functions of sound in relation to the game form *Audio Games*, in order to explore the potential of sound when acting as an autonomous narrative form. Because this is still a relatively unexplored research field, the main purpose of this study is to help establish a theoretical ground and stimulate further research within the field of audio games. By adopting an interdisciplinary approach to the topic, this research relies on theoretical studies, examinations of audio games and contact with the audio game community. In order to reveal the roles of sound, the gathered data is analyzed according to both a contextual and a functional perspective.

The research shows that a distinction between the terms 'function' and 'role' is important when analyzing sound in digital games. The analysis therefore results in the identification of two analytical levels that help define the functions and roles of an entity within a social context, named the *Functional* and the *Interfunctional* levels. In addition to successfully identifying three main roles of sound within audio games—each describing the relationship between sound and the entities *game system*, *player* and *virtual environment*—many other issues are also addressed. Consequently, and in accordance with its purpose, this study provides a broad foundation for further research of sound in both audio games and video games.

Keywords:

Audio game machinery, Sound analysis, Game studies, Roles of sound, Virtual environment, Accessibility, Sound perception, Auditory worlds, Sound functionalities, Transdiegetic communication

Acknowledgements

To my thesis supervisor Johnny Wingstedt
for all your great guiding and many inspirational talks,

To my dachshund Vid
for reminding me to take long, thoughtful walks once in a while,

And to my girlfriend and editor, my lovely linguist Lucija Antolković
without whom this thesis could never have been written,

...THANK YOU FOR ALL YOUR SUPPORT!

– RICKARD ÅSÉN

Table of Content

1 Introduction

Originally intended as interactive entertainment for people with visual impairments, the *Audio Game* is a rare form of digital game production that provides a full gameplay experience through aural means alone. Although the term 'audio games' is still an unknown concept to most people, recent developments have increased the popularity of the game form among both visually impaired and sighted players. The first time I heard the term, I was somewhat puzzled by its meaning—but intrigued by the idea. As a student of Sound and Music Production, with an interest in digital gaming, the concept seemed to fit my interests very well. Exploring the game form further, I found that although a small scene of audio game developers exists, the research field appears to lack any major theoretical conventions. Compared to the field of sound in video games—which in itself is highly underexplored—the research on sound in audio games is infinitesimal. Although studies have been done on the functionalities of sound in various audio–visual media, the role of sound when acting as an autonomous narrative form, in an interactive virtual environment, has not yet been explored.

1.1 Definition of Function and Role

During the research of this thesis, it has become evident that the terms *function* and *role* are often mixed up and used synonymously in academic literature. Although similar by definition, the terms are in fact slightly different in meaning. Oxford Dictionaries defines the two words as follows:[1]

> **Function**
> "An activity that is natural to or the purpose of a person or thing"
>
> **Role**
> "The function assumed or part played by a person or thing in a particular situation"

To summarize the two: when a person or thing acts as a role, it performs functional activities within a context. A function is therefore an activity or purpose assigned, and subordinate, to a role. Simultaneously, however, a role is often defined by its functions. It is according to this interpretation, that this research aims to explain the *role* of sound in audio games by studying both its *functions* as well as its *context*.

[1] http://oxforddictionaries.com/ See bibliography ("Function," n.d.; "Role," n.d.).

1.2 Purpose of the Study

The purpose of this study is to help establish a theoretical ground and stimulate further research in the field of audio games. By analyzing the concept of audio games from several angles, and exploring the various functions related to sound within these games, the aim of this thesis is to reveal and describe the role of sound when acting as an autonomous narrative form within an interactive virtual environment.

1.3 Research Questions

- What is the role of sound in audio games?
- How does sound function within audio games?
- In which ways are the functionalities of sound in digital games affected by lack of visual stimuli?

1.4 Demarcations

A comprehensive study of all aspects of the audio game phenomena is not possible within the frames of this thesis. Therefore, this thesis does not attempt to create new theoretical conventions to the field of audio games as a whole. It does, however, aim to study audio games from several angles in hope to help future studies form a more comprehensive overview of the field of audio games. Although this thesis will discuss audio games as a concept, the main focus of the study will be to identify the *role* of sound within the game form.

More specifically, this thesis will focus on sound in first-person three-dimensional (hereafter referred to as *F-P 3-D)* audio games,[2] even though many of the findings can be applicable to other types of audio games as well. In addition, although music is a form of sound commonly present in digital games, its uses and functions in audio games will not be explicitly examined in this study.

1.5 General Terminology Used in the Thesis

Many terms, words and concepts that will be used throughout this thesis may not be common in the vocabulary of many readers. In an effort to make this thesis more comprehensible, most of this terminology will be explained during the text. However, a prior knowledge of sound or game design might be helpful when reading this thesis.

[2] First-person three-dimensional (or F-P 3-D) games refer to the type of games that uses the perceptual perspective of the avatar to describe a three-dimensional diegetic environment.

1.5.1 Stipulations of the Digital Game Forms Used in This Thesis

Although audio games are often referred to as a *genre*, this thesis will instead refer to audio games as a *game form*.[3] This relates to the distinction made in this thesis between *Audio Games* and *Video Games*—both of which are subcategories of the overarching term *Digital Games*. In this thesis, the term 'audio games' will refer to the digital type of game that reproduces the diegetic and extradiegetic environment primarily based on aural content.[4] Whenever a game represents the diegetic environment with iconic or non-arbitrary visual graphics, regardless of the amount of aural content, this thesis will instead use the term 'video game'. This thesis also stipulates to use the term *Audio-mostly Games* for certain audio-based games, which includes enough visual graphics to change the functionality of the game if played by a sighted player instead of a visually impaired player. Audio-mostly games can thereby be seen as a game form in between audio games and video games. These games are fully playable by the sense of hearing alone, but they often represent the diegetic environment with extradiegetic arbitrary graphics (e.g. maps of the environment), or they present other graphical elements that in one way or the other affects how the games can be played.

1.5.2 Stipulations of the Three Levels of Digital Game Development

This thesis will on many occasions refer to three different levels of digital game development: *hobbyist*, *indie* and *mainstream*. The hobbyist level refers to individual persons or very small game development teams, which produces games on low or non-existing budgets and with low commercial interests. These games are most often shared as freeware among hobbyist game communities. The term 'hobbyist', however, does not refer to the qualifications of the developers, as these may be anything from novices to experts in game development. The indie level refers to small development teams with somewhat larger commercial interests. Though still working on relatively small budgets, these teams produce games with the intention of creating a product available for sale to customers through various channels. The indie-based teams work independently from the larger publishing companies. This gives the indie-companies much creative freedom, which often results in more experimental games. The mainstream level consists of large development teams, financed and supported by large

[3] The term 'genre' is misleading in this case as there are many different genres within the game form audio games.

[4] The concept of diegesis will be further explained later, in segment *3.3.3 – Diegesis in Audio Games.*

game publishing companies. These game productions have the largest budgets, but also the most creative constrictions as their high production costs requires the games to be sellable to a larger market of players. There are of course variations of these levels, but this thesis is written based on these general explanations. As will be discussed in more detail later, audio games are still in a relatively experimental game development phase, and are therefore primarily produced on a hobbyist level, with some companies venturing into audio game production on an indie level.

1.6 On My Perspective as a Writer

In the art world of audio games today, the blind are the norm and the sighted a minority. With this in mind, it is important to mention that I am myself a fully sighted person, and thus this thesis is written from that perspective. Through various discussions on the internet forum forum.audiogames.net, I have come to realise that the majority of audio games produced are developed as a hobby project, often created by visually impaired programmers. Professional game development teams such as Somethin' Else—responsible for titles such as Papa Sangre and The Nightjar—are thus also a minority in this art world. As a student of Sound and Music Production, my aims are set for a career in the professional game industry. I am not a programmer myself, but I consider myself to have enough experience with computers and digital games to be able to appreciate the complexity and technical limitations involved in developing a game. In addition, my gaming experience is primarily based on playing mainstream video game titles.

By saying this, I want to emphasize that this thesis is written from a perspective of a sighted player with limited gaming experience of audio games, but with a background in sound production and an extensive experience in playing mainstream video games. Since previous research is still scarce in this field, I believe that further research from other perspectives is needed in order to create a truly comprehensive view of the concept of audio games.

2 Method of Research

This thesis has focused on a theoretical interdisciplinary approach to audio games. The empirical data that is used in the study has been gathered and categorised based on two different analytical viewpoints on the field of audio games. The first viewpoint, examined in *Chapter 3 – The Conditions of the Audio Game Machinery*, focused on understanding the underlying conditions of audio games, and how these relate to the aspects of development, game system, player and sounds within the virtual environment. The second viewpoint, examined in *Chapter 4 – The Various Functions of Sound in Audio Games*, has instead focused on outlining the different functions of sound within audio games, in order to more closely define how and why sound is utilized in the games, and describe the functional interrelationship between sound and player. By analyzing and comparing the findings of these two viewpoints in *Chapter 5 – Summary: The Role of Sound in Audio Games*, the study aims to define what the role of sound in audio games is—and why.

2.1 Information Gathering

In order to form this thesis, a theoretical interdisciplinary approach was adopted, including literary study of research from many fields, contact with the audio game community, examination of audio games and information gathering from various internet sources.

2.1.1 Literary Study of Related Research

The main method used in this study was to adopt a literary approach, and draw information from a wide range of related research fields. This section will briefly present the main literature and research studies used in this thesis.

The book that initially inspired this research was Kristine Jørgensen's *A Comprehensive Study of Sound in Computer Games: How Audio Affects Player Action* (2009). In her research, Jørgensen presents a detailed theory on how sound functions in computer games, based on a series of listening tests and interviews with players in combination with a rich theoretical background. The second source on game audio functions is called the *IEZA*-model, developed by Sander Huiberts and Richard van Tol (2008; 2010). In short, the IEZA-model is an analytical framework for categorizing sound events in digital games. Besides research on game audio, this thesis has also

studied the research on film sound listening from the book *Audio-Vision* by Michel Chion (1994), as well as Alan Beck's (1998) studies of sound usages in radio dramas. In addition, Michael Forrester's (2002) study of sound imagery was primarily used in order to cover the psychological aspects of sound perception. To understand communicative functions on a meta-level, studies have been done on Michael Halliday's (1978) research in the field of systemic functional linguistics, as well as the subsequent work on multimodality researched by Theo van Leeuwen (1999) and Johnny Wingstedt (2008). And lastly, throughout this research, Murray R. Schafer's (1994/1978) seminal work on acoustic ecology, *The Soundscape*, has provided much inspiration and understanding of sound on a more philosophical level.

2.1.2 Contact with the Audio Game Community

The largest social network of the audio game community is the website www.audiogames.net, originally developed by the previously mentioned IEZA-model creators Richard van Tol and Sander Huiberts. On this website exists a very active forum, where audio game developers and players discusses various aspects of audio games. Throughout this research, this forum has been a great source of information on audio game related aspects, and has been used frequently to steer this research in the right direction. At one point, I created a forum thread in which I asked six questions related to this research, to which I immediately received many useful answers from the members.[5] This was initially intended to be a repeating method of collection data, however, as the form of the thesis developed, I had to prioritize other methods of information gathering instead. That forum thread, however, as well as several other threads on that forum, was very helpful in writing section *3.1 – Conditions of the Game System & the Development Scene* of this thesis.

2.1.3 Examination of Various Audio Games

During this research, numerous audio games have been played or in other ways examined by myself. A few of these have then been selected to exemplify certain sound functionalities in audio games. Although many of these exemplified games have been played during the study (e.g. *BlindSide*, *Swamp*, *Zombies, Run!*), some games have been unavailable for me to play personally (e.g. *Papa Sangre*, *Deep Sea*, *Tim's Journey*). In all cases, I have gathered additional information from a variety of sources—if not from

[5] The forum thread is available at http://forum.audiogames.net/viewtopic.php?id=9410.

the games themselves, then from audio-visual clips of the games being played, articles, reviews, the www.audiogames.net forum or the official websites of the games.

Most of the exemplified games have been developed by the professional indie-game scene, which, as stated in section *1.6*, may not be representative for the whole audio game development scene. However, considering this thesis' focus on F-P 3-D audio games, the games selected will hopefully provide adequate representation of that demarcation.

2.1.4 Information Gathering From Various Internet Sources

In addition to the other methods, a considerable amount of time has been dedicated to an extensive search for information from various internet sources. Although this may relate closest to the *3.1 – Conditions of the Game System & the Development Scene* section, it has been used throughout the thesis to gather information, discover new information, double-check old information or search for opposing views of both new and old information.

3 Conditions of the Audio Game Machinery

In order to understand the role of sound within audio games, it is important to first understand the sound within its context. This chapter will therefore provide a brief introduction on the current status in the audio game development scene. In addition, this chapter will look at the conditions at work when playing an audio game, examining game system and player aspects, as well as the sounds themselves.

In this thesis, a digital game will be considered as one body—a machine—comprised of three separate entities: the *game system*, the *player*, and the *virtual environment*. Each of these serves a vital purpose in the machinery. By removing the player from the equation, the game machinery is left with a virtual environment without actions, interactions or audience. By removing the game system or parts of it, the virtual environment would cease to exist or hinder the player from interacting with it. Finally, by removing the virtual environment, neither game system nor player has purpose. Accordingly, this chapter is divided into three main sections, discussing each of the entities of the game machinery separately. However, since the three entities interrelate at all times, so will also these sections at some points.

3.1 Conditions of the Game System & the Development Scene

This first section of the chapter will act as an introduction to the audio games themselves. This section will discuss the current state of the audio game development scene, and highlight vital technological aspects involved in creating an audio game. Since very little research has previously been done on the audio game scene, the data used to write this section is primarily based on observations of the audio game scene, various internet sources and examination of the audio games themselves.

3.1.1 Definition of the Game System

All digital games need platforms. Whether this is a game console, an arcade machine, a computer or a mobile phone, the technological platform is a prerequisite for all forms of digital games. In her study of sound in computer games, Jørgensen (2009) uses the research field of auditory display and labels this technological aspect of digital games as the *User System* (p. 55). Jørgensen describes her research as a "study of the relationship between the *game world* as a coherent environment in which the player acts, and the game as a *user system* that is manipulated through an interface" (pp. 3–4).

The term ‘user system’, however, can easily be confused with the term ‘user–system interaction’, which is common within the research field of human–computer interaction and the field of human factors. Sikorski (2003) writes that “[u]ser–system interaction […] relies on investigating the interaction between the user and the system in the specific context of use” (p. 411). Accordingly, ‘user–system’ in that sense refers to the study of the interaction *between* a system and its user, rather than a system *for* the user, as Jørgensen’s ‘user system’ implies. I do not disregard Jørgensen’s term as a valid definition of the technological system used by games. However, I believe that ‘user system’ might direct added attention to the relationship between the player and the system, rather than defining the system as a separate entity in the game machinery—equivalent to both *Player* and *Virtual Environment*—which is the approach assumed in this thesis.

To avoid confusion, the term *Game System* will be used in this thesis to describe the technological link between the in-game environment and the player. Despite being commonly used in the research field of digital games, it is difficult to find any previously detailed descriptions of the term ‘game system’ in the research literature. Therefore, the usage of the term ‘game system’ in this thesis hereby stipulates to include: the complex software of underlying programming code that defines the fundamental rule system of the game; the performance inducing hardware that constitutes the computing platform (e.g. a game console, an arcade machine, a computer, a mobile phone); and the physical equipment and assisting tools directly used by the player (e.g. keyboard, mouse, joystick, touchscreen, headphones, visual display). Being situated externally from both the in-game world environment and the player’s cognition, the game system provides the essential bridge of communication between the two. It is through the game system that the player is provided with the information present within the virtual game world. In return, the player utilizes the game system to communicate commands back to the game world, in order to control the avatar or other game elements, and thus create a progression of events in the game.

3.1.2 Definition of Gameplay

Within digital game terminology, the interaction between the player and the game is commonly referred to as the *gameplay*. Björk & Holopainen (2004) define gameplay “simply as the structures of player interaction with the game system and with the other players in the game”; further on, they claim that gameplay therefore “includes the

possibilities, results, and the reasons for the players to interact within the game" (p. 3). Since these terms, 'game system' and 'gameplay', are general terms for all digital games, their usage within the field of audio games can be assumed to work in a similar manner as when used in relation to video games.

In many video game magazines,[6] the gameplay of a game is often mentioned and rated as a separate category in their review system, in addition to other aspects of the games. This conceptualizes gameplay as a subjective notion that can be quantifiable on a scale from 'bad' to 'good'. However, to define what constitutes the gameplay of any given game can often be difficult to specify, and what might be considered 'good' gameplay qualities in one game might not apply in another. As seen in the description above, the concept of gameplay can be said to relate to the degree of enjoyment the player experiences based on how the control system and challenges are designed in the game, and how well the communication functions between the player and the game.

Being a subjective notion, the player's view on the gameplay of any given game is highly related to the previous gaming experience of that player. This makes an interesting point in the study of audio games, as it is an unconventional form of gaming for most people. For sighted players, the gameplay experience is directly affected by the removal of the player's sense of vision. Left with a blind perception of the virtual environment, new players often face a steep learning curve, as the level of difficulty is increased not only by the unfamiliar game, but also by the unfamiliar perception mode.

3.1.3 Developing Audio Games – A Scene under Development

Having never been a part of the mainstream digital game scene, audio games have instead mostly been the subject of experimental research in the development of entertainment for the visually impaired, or a non-profitable project to the hobbyist game developer—often programmers that are visually impaired themselves.[7] Up until now, audio games have predominantly been developed for the PC computer, running any version of the Windows operating system. Recently however, as technology has made it easier to develop games for other platforms, a small number of audio games have started to appear on the Mac OS X and the mobile phone applications market as well.

[6] E.g. Gamereactor (www.gamereactor.eu), Gametrailers (www.gametrailers.com) and 148apps (www.148apps.com)

[7] The largest community of audio game developers is www.audiogames.net, where the majority of users are visually impaired to some degree.

The advancements in the mobile phone industry during the last few years have enabled small game developer teams to produce and release digital games more easily than ever before. By selling games as downloadable purchases to mobile phones, a scene of professional audio game developers is slowly starting to arise. However, judging from discussions on the audiogames.net forum, this scene does not seem to have evolved from the previous hobbyist producers, but rather from the more commercially established independent video game developer scene (CAE_Jones, Appendix B.1).

One of the most renowned games of this indie-scene is the F-P 3-D thriller game Papa Sangre (2011). As is common in many other audio games, Papa Sangre uses the impact of sense deprivation to emphasize horror elements within the game. The gameplay approach in this game is fairly simple. The player uses the touchscreen of an iPhone to rotate the avatar left or right by swiping the finger on the top half of the screen. By tapping the lower half of the screen, the player then moves the avatar forward in the direction it is facing. The challenges of the game are then related to interaction with the game environment, by for example following and collecting certain sound cues or avoiding potentially dangerous objects or monsters. The easy-to-use control system is combined with an interesting game plot and a well-produced sound design.

One of the best-selling audio games for the mobile phone market is called Zombies, Run! (2012). This game engages the player to physically jog in the real world, while the gameplay and the story are being unfolded through the player's earphones. As such, this game differs a lot from the more conventional audio games, since it involves certain physical efforts from the player in order for the game to progress. The fact that this sound-based game is released on a portable medium allows it to be a very mobile experience, where the player is provided with all the information necessary for the gameplay through sound, while being able to engage in other physical activities at the same time. The interaction between the player and the virtual environment is, however, rather limited in this game, and the story is presented similarly to interactive audio books. Nevertheless, it does have certain gameplay assets that would qualify it as a game. Zombies, Run! (2012) is often called an *ultra-immersive* game, as it integrates the real world of the player with the game world environment, directly immersing the physical player into the narration. By utilizing the built-in Accelerometer or GPS functions of the phone, the game measures the location and speed of the player, involving the physical actions of the player as part of the game system, and links these actions to events in the game world.

According to statistics from Appshopper.com, Zombies, Run! is, at the moment of writing, ranked #16 on most-sold-application in the Healthcare & Fitness section ("Top 200," 2012), and further details show that it has at one or many occasions been ranked as #1 on the same list. Due to privacy policies, however, it is difficult to retrieve exact figures of revenue from purchased apps, which makes it problematic to form a definite list on the success of various games. Nevertheless, it is possible that Zombies, Run! is the most economically successful audio game to date. The key ingredient to the game's success is most likely its approach as being a combination of fitness application and ultra-immersive audio game, with the twist of a zombie-survival story. This game however, is very far from the current norm of audio games. Developed independently from the hobbyist scene, Zombies, Run! seems to target the mainstream audience rather than the visually impaired community. Even though the actual gameplay is audio-based, the menu interface is instead very graphical. In addition, the ultra-immersive approach integrates the visual world of the player's environment with the aural virtual world of the avatar. In this case, it might therefore be more correct to label the game as an *audio-mostly* game, since the game is—technically speaking—playable by hearing alone, but the player's sight may enhance the gameplay experience and possibilities. To make the game more accessible to visually impaired players, the developers has implemented support for the Apple developed screen reader VoiceOver,[8] which can be used to read the interface text in Zombies, Run! (Alexander, 2012; Watson, 2012). In addition, when utilizing the phone's accelerometer, the game can be played while running on a treadmill; a very important function for visually impaired players. When running outdoors, the player needs to rely on the four other senses to percept the real world environment, since the game requires him or her to dedicate their hearing to the virtual game world. If the player then already relies heavily on hearing to navigate the real world, it would make jogging in a natural environment extremely difficult if that sense is blocked as well. Or as one visually impaired member of the audiogames.net forum says, "Well I would like to run, but running as a blind guy in a city full of traphic [*sic*] while wearing headphones, no way!" (SLJ, Appendix B.2).

[8] VoiceOver is an accessibility application that reads textual content on the screen out loud. For more information, see: http://www.apple.com/accessibility/iphone/vision.html

3.1.4 Turning Audio into Sound

The audio game Papa Sangre (2011) uses a custom-built audio engine that, according to the developer's official website, is the first ever audio engine to utilize real-time binaural processing of game audio on a handheld device ("About," n.d.; "First Post!," 2010). This engine uses so-called *HRTF*-processing to simulate the binaural hearing process normally experienced by humans. *HRTF* is an abbreviation of the term 'Head-Related Transfer Function', which relates to the way sound travels around the human head. Møller, Sørensen, Hammershøi & Jensen (1995) write that a HRTF is "a transfer function that, for a certain angle of incidence, describes the sound transmission from a free field to a point in the ear canal of a human subject" (p. 300). Even though the technology exists today, besides Papa Sangre, it seems nearly impossible to find any other commercial game—video game or otherwise—that uses a somewhat realistic real-time binaural audio processing. While many modern mainstream video games are mixed for 5.1 or even 7.1 surround sound systems,[9] not one of them seems to incorporate the use of real-time binaural stereo reproduction. It appears that the mainstream video game industry has not yet prioritized to adopt the possibilities of binaural sound, even though many players are using headphones instead of a surround sound speaker system, especially when playing PC games. There are, however, a few workarounds to this issue.

In the last decade, there has been an increase in the development and sales of multi-channel surround sound headphones, in which multiple drivers are implemented within each earpiece. The quality of sound reproduction through such headphones has, however, been under much debate.[10] Another solution is to use stereo headphones and implement binaural simulations through a third-party developer's software. This software technology processes the multi-channel audio output of the games, and applies binaural algorithm filters to simulate a HRTF-effect. One of the major differences between using an audio engine with real-time binaural processing directly in the game, compared to using third-party middleware, is that the audio engine calculates the exact position of the sound directly in the game, whereas the middleware has to reverse engineer the effect based on audio from 5 or 7 different channels. An issue to consider

[9] E.g. Killzone 3 (2011), Batman: Arkham Asylum (2009), Warhawk (2007).

[10] There are a wide variety of threads in various forums that discusses this issue. For example: http://www.head-fi.org/t/593050/the-nameless-guide-to-pc-gaming-audio-with-binaural-headphone-surround-sound/1020#post_8936686

when using any kinds of HRTF-algorithms, however, is that all human heads and ears are shaped differently, and thus the parameters used might not be accurate to accommodate the long-term habits of hearing patterns developed by the player.

It is not the aim of this thesis to further research which of these technologies is best applied for reproducing surround sound through headphones. However, what seems to be the general opinion among audiophiles is that the best sound quality for the money spent comes from using a good pair of stereo headphones with either of the binaural processing principles. Headphones with multiple drivers are costly to produce, so in order to sell them in a price range affordable to most gamers, it would make sense that the quality of each driver is considerably inferior to drivers in stereo headphones in the same price range. In addition, multiple-channel headphones also have to account for the physical limitations of the earpiece. In order to fit 3 or 4 drivers within one earpiece, they would have to be notably smaller than drivers of stereo headphones, which usually leads to poorer bass reproduction. This is, however, not an issue that has been tested in this research, and judging from the difficulties in finding research on the comparison of these three principles, further research on the topic would be welcome.

In games based solely on sound, the reproduction of an accurate soundscape is crucial for the gameplay factor. As can be seen above, optimization of the game system (including platform, hardware assets, audio engine, headphones / loudspeakers, etc.) is an important issue when managing how the various aural resources are performing, and how the soundscape will be perceived. Although real-time binaural-processing can be quite demanding on the game system—which may explain why so few games have utilized it so far—recent developments, such as the binaural audio engine developed for Papa Sangre (2011), show that the technology may now be mature enough to enable the use of real-time binaural-processing in more digital games.

3.1.5 Budget – Making Ends Meet

One of the key issues in the development of an audio game is the budget. As mentioned previously, audio games have never played any major part in the mainstream digital games market. Even the most successful audio games cannot compare to the budget and profits of leading mainstream video games. This has been a major factor in how audio games has been developed so far. The majority of audio games have been, and are still, developed as hobby projects with little or no hope of gaining any commercial success. Thus, they are most often developed by a single person, or perhaps by a small

team of 2–3 people, as a side project to their other occupations. In addition, considering the technological difficulties involved in creating any kind of game, most of the audio game developers on a hobbyist level seem to have a background in computer programming rather than sound production.

With no budget to speak of on the hobbyist level, the possibilities of acquiring professional audio equipment, sound designers and voice actors are very slim. Creating a realistic three-dimensional audio environment, with hours of recorded audio events and dialogue to fill it with, is therefore an immensely difficult task for the lone programmer. Adding to that the technological limitations that still exist with the easily available audio engines of today, it is nearly impossible for a small team with no budget to realise a fully functional and satisfactory three-dimensional audio game. Instead, most of the hobbyist developed audio games are more simple, two-dimensional games.[11] There are of course exceptions, and a few developers from the hobby scene have created three-dimensional games that gained popularity within the audio game community. Most of these being freeware, however, the creators have received little or no financial gain from these projects. Despite having received popularity, even these games could most likely have been improved in terms of functionality, regular patch maintenance and sound design if they had also had the possibility of financial sustainability.

Judging by the current status of the audio game market, one can see clear tendencies as to which platform seems most suited for financial gain when developing audio games. For various reasons, the hobbyist game designers have developed most of their games intended for the PC platform. These games are easy to share among the community, and the technical requirements of these games are often the easiest to fulfil by the largest amount of intended players. The indie game developers, on the other hand, have focused mostly on the mobile phone market. The most common channels for selling or buying games and applications on the mobile phone platform are the various application stores.[12] By utilizing these channels, the developers are able to fairly easily put a price on their product and receive money for each purchase. This market, however, is somewhat self-regulatory, in terms that if the game is not developed well enough, fewer people will buy it and the developers will receive less financial gain.

[11] This does not mean that these games are any less enjoyable, and many of these two-dimensional games are immensely popular on the www.audiogames.net forum.

[12] Such as *Google Play* for Android phones and the *App Store* for Apple devices.

The developers therefore have to walk a fine line regarding how much money and effort to put into the project of creating a game. If a game is not produced well enough, chances are that it will sell badly. On the other hand, too much time and effort spent on the project might lead to an abundance in expenses that the income of the final product might still fail to meet, since the targeted group of buyers is too small. Developing games for a specific platform, such as an iPhone, gives the developer more control and knowledge of how well the game will actually function when used by the end-consumer, as opposed to developing to the PC platform, where the technical setup might vary greatly between one consumer and another. It does, however, also narrow down the target group further if the platform is too specified. An audio game developed exclusively for the iPhone will for example exclude all users of Android or other phones.

While making sure a game reaches the targeted market group and starts generating revenue to its production company is difficult, perhaps the most difficult part in developing an audio game would be to initially gain enough financial support to get started on a project. This is of course true for all media productions, but since the art world of audio game development is still very young and untried, it could be particularly difficult to predict whether or not an audio game is a potential success or a fail. The recent development, however, with indie-based professionally developed audio games such as Papa Sangre (2011) and Zombies, Run (2012), might indicate a brighter future for the more professional audio game development scene, which in turn may generate a positive effect on the hobbyist level as well.

3.2 Conditions of the Player

What separates digital games from other, non-interactive, audio–visual media is the presence of a *player*. As discussed in the beginning of this chapter, the player is a vital entity in the digital game machinery. Within the machinery, it can be argued that the player acts an ambiguous role, having both a passive and an active approach simultaneously. In its passive approach, the player acts as an observer of the virtual environment and an audience to the game narrative—a recipient of a pre-designed story. In its active approach, on the other hand, the player interacts with the virtual environment via the game system, dynamically altering the progression of the game narrative. These two approaches are delicately intertwined and support each other at all times throughout the gameplay experience. This section will therefore discuss some of the mental processes involved when a player observes and interacts with an audio only environment, as well as explore the player's relationship to sound in audio games.

3.2.1 The Player–Avatar Relationship

In order for the player to experience the virtual environment, the game system must constantly provide information about the game world to the player through various means. When observing the virtual environment in a first-person game, the player uses the perspective of the avatar to perceive the in-game world.[13] Like a puppet on a string, it is the avatar that directly interrelates with the virtual environment, not the player. The avatar can therefore be said to function as a channelling medium—an extension of the player's ears and eyes in the virtual environment—with the game system working as a transducer of information between the digital dimension and the real world. The purpose of this connection is to transfer the simulated sense impressions of the avatar to the player, in order for him or her to make an informed decision on how to further interact in the game. Since computers are generally designed to present their digital content through visual and aural modes of expression, the sight and hearing are the only senses commonly represented in digital games today. There are, however, many digital games that also utilize haptic technologies—such as a vibration generator in the game controller—to generate a certain degree of tactile feedback (e.g. an extra rumble effect at certain points in the game). This tactile effect, however, does not depict more complicated sense impressions of the avatar, such as temperature, a gentle touch or a

[13] The avatar is the game character controlled by the player—most often the protagonist of the story.

mild wind in the hair, nor are there currently any digital games that have implemented the remaining two human senses: smell or taste.

In audio games, where the visual is omitted fully or partially, the essential information used to describe the virtual environment is communicated through aural representation. This limitation sometimes leads to games being created to involve the lack of sight within the game plot, in order to explain to the player why he or she cannot see anything. The avatar might for example be blind, or the game plot is set in an environment where there are no visual stimuli to be found, such as in Papa Sangre (2011) where the avatar has been transferred to a 'pitch-dark' spiritual realm. However, these explanatory plots are not a rule in audio games, and many games either do not explain the lack of visuals to the player at all, or the avatar might even be sighted but the game still only conveys the aural information to the player. In the game BlindSide (2012), the developers use a radio drama technique called *interiorizing*, where the avatar either speaks its thoughts out loud, or the player (listener) is positioned inside the avatar's mind and is thus able to hear its thoughts. In his study on radio dramas, Alan Beck (1998) further explains the concept:

> "[Interiorizing] is a powerful way for the playwright to let the listener into the internal dialogue of the character's 'me' within the outer 'I'. I call this production convention 'interiorizing' and it is where radio scores over plays in other media. On stage, the internal monologue has to be externalised and the actor is seen to deliver the lines. I regard 'interiorizing' as radio's 'fourth dimension' and it can establish a complicity with the listener, a process as familiar as our own inner ruminations."
>
> (Beck, para. 10.7)

In BlindSide (2012), the thoughts of the avatar function as a guide to the player, and are often used to describe objects in its surroundings. This assists the player in getting a mental overview of the immediate environment, and helps with certain gameplay challenges.

3.2.2 The Difference Between Listening and Hearing

Whenever a sound is heard, it is important to define the relationship between the listener and the sound, in order to understand how the listener approaches that sound. To begin

with, there is the matter of hearing or listening. On this, Jørgensen (2009) writes that "[t]here is a perceptual difference between these in the sense that *hearing* is an unintentional activity while *listening* is intentional and focused towards specific sounds" (p. 73). The human brain is very good at selectively choosing what to listen for in a sea of sound—a phenomenon commonly referred to as the *cocktail party effect*[14]—and this applies to all listening situations, in natural or virtual environments alike. This selective listening is an interesting notion to keep in mind when studying the player experience of audio games, where the gameplay experience is highly dependent on the player's ability to discern what is happening within the auditory environment.

In an attempt to outline the listening processes active while playing computer games, Jørgensen (2009) combines the research of film sound by Michel Chion (1994) with the study of music listening by Denis Smalley (1996). Jørgensen summarizes four different listening mode situations relevant to game audio (p. 78):[15]

1.) Causal Listening (Chion)

> When the player is actively listening for the source of the sound, to gain information about its cause.

2.) Semantic Listening (Chion)

> When the player is actively listening for the content of the source, attempting to gain a conscious understanding of the semantic meaning.

3.) Reduced Listening (Chion)

> When the player is actively listening for properties and traits of the sound itself.

4.) Reflexive Relationship (Smalley)

> When the player passively hears a sound and only responds on an emotional level to it.

Jørgensen motivates her summary by arguing that although Chion's theories are easy to understand and apply well to video games, they do not provide a mode for passive hearing or the emotional responses to sound, whereas Smalley's theories do (p. 77).

[14] Termed by the ability of people to have a conversation in the middle of a crowded room.

[15] Jørgensen presents the ideas of Chion and Smalley in more detail, but this study will only present the ideas most suitably adapted for audio games.

These listed listening modes active while playing video game can be applied when playing audio games as well. What differs between the two game forms is instead the way the listening modes function. In addition to his three listening modes, Chion (1994) describes a situation called *Acousmatic Listening*, in which the listener hears a sound but does not see its source. In relation to *Causal Listening*, Chion adds that "[w]hen the cause is visible, sound can provide supplementary information about it" whereas "[w]hen we cannot see the sound's cause, sound can constitute our principal source of information about it" (pp. 25–26). Since sounds depict events and not objects,[16] the listener in an acousmatic situation will attempt to ask for information about the source—the object. Chion writes that the acousmatic situation will intensify causal listening, since the aid of sight is removed (p. 32). Chion further explains:

> "Confronted with a sound from a loudspeaker that is presenting itself without a visual calling card, the listener is led all the more intently to ask, "What's that?" (i.e., "What is causing this sound?") and to be attuned to the minutest clues (often interpreted wrong anyway) that might help to identify the cause."
>
> (Chion, 1994, p. 32)

Since audio games by their nature present the environment through an acousmatic situation,[17] the listener therefore has to actively seek to identify the cause of every sound, and use the sound as the primary source of information about the object or event. In video games, on the other hand, the source is most often visible and known to the player, and consequently, sound only reveals additional information about its cause.

Chion (1994) also identifies the term *verbocentric*, which refers to the human tendency to automatically focus the attention to voices, rather than other sounds. He explains:

> "When in any given sound environment you hear voices, those voices capture and focus your attention before any other sound (wind blowing, music, traffic). Only afterward, if you know very well who is speaking and what they're talking about, might you turn your attention from the voices to the rest of the sounds you hear. So if these

[16] A concept that will be further discussed later, in segment *3.3.1 – Sound As Events.*

[17] As there is no visual representation of the virtual environment in audio games.

> voices speak in an accessible language, you will first seek the meaning of the words, moving on to interpret the other sounds only when your interest in meaning has been satisfied."
>
> (Chion, 1994, p. 6)

In an acousmatic situation, this could mean that listening to semantic use of voice would take precedence over causal identification of an object. Consider, for example, that an unknown sound event is being played, and in relation to this a voice describes the cause of the sound. If the description relates close enough to the properties of the sound, the listener would probably believe the voice—even though the listener might have evaluated the cause differently if the voice had not been present. This could be particularly true in audio games, since the player is aware that the sounds are just representational recordings, used to describe a virtual environment. The player might therefore accept a discrepancy between the represented sound and its natural counterpart, as long as it is believable.

If the player is unable to determine causally what the sound represents, the player instead needs to listen for properties and traits in the sound—utilizing the reduced listening mode. If abstract sounds are used within an audio-only production, such as synthesised sounds with no natural counterpart, it is therefore important to create an *agreement* with the player, informing him or her what the sounds represent. Otherwise, the player will not be able to objectify the sound event[18]—possibly creating a confusing soundscape. When this agreement exists, and the player is aware of what the sound represents, the player can return to utilize the causal listening mode instead. In video games, on the other hand, synthesised sounds can be used more freely, since the represented source of the sound event is often visible and easy to objectify by the player. The reduced listening mode can therefore also be argued to relate to the listeners ability to recognize and objectify a sound event. However, reduced listening is still most commonly utilized when a sound event is already causally determined.

Jørgensen (2009) writes that computer games very rarely utilize reduced listening (p. 78). Although this is mostly true for audio games as well, implementing challenges based on the player's use of reduced listening could lead to some interesting audio game concepts. For example, it could perhaps be utilized in an audio puzzle game, or in an

[18] The *objectification* process will be discussed in the next segment, *3.2.3 – The Sound Stage of Audio Games*.

abstract audio adventure game that utilizes purely synthetic sounds without reference to real world objects. In the latter, the gameplay challenges could rely on the player's ability to listen intently on the qualities of the sounds themselves, and perhaps to modify these in order to progress the game. This is purely theoretical though, and might be difficult to apply into an enjoyable game.

3.2.3 The Sound Stage of Audio Games

Drawing from the studies of radio dramas—an art form that could be considered having an equivalent relationship with audio games as film has with video games[19]—enables us to think of a listener perspective in relation to the virtual environment. In film studies, a common concept is the *point-of-view,* which encompasses the viewer's perspective to what is seen on the screen. Expanding this notion to the auditory side of media productions, Chion (1994) constructs the similar concept of *point-of-audition*—to explain the listener's perspective on what is heard (pp. 89–94). Beck (1998) further explores this concept of listener perspective by identifying the notion of the *Sound Frame*. The sound frame is an allegory to film's concept of *On-screen* vs. *Off-screen.*[20] On studying the work of Beck, Forrester (2002) writes the following:

> "Producing sound events that are 'in-frame' (rather than out-frame) involves balancing together the sounds with explanations for their occurrence within a 'sound perspective' based on what we would normally expect with our everyday acoustic environment, i.e. sounds close by us. In contrast 'out-frame' sounds are produced as if external to the immediate present (for the participants in the drama), e.g. the sound of a dog barking in the distance, seagulls and so on."
>
> (Forrester, p. 44)

If applied to F-P 3-D audio games, the concept of sound frame would imply that the game developers are able to create a sort of auditory frame, within which the virtual environment and gameplay content is presented to the player through layers of varying degrees of importance. Through selective listening, the player can then 'zoom-in' on any sound of interest, with the sound frame present to guide and propose sounds for the player to listen to. A well-designed sound frame would therefore function as a support

[19] Radio dramas and films both present a virtual environment but lacks interactivity.

[20] A term used to define events occurring either within or outside the picture frame.

for the player to visualise the immediate situation of the avatar. Adding the interactivity, however, enables the player to actively interact with the surroundings—as opposed to being a passive receiver of a pre-mixed soundscape, as is the case with radio dramas. Thus, the sound frame perspective is continuously undergoing a metamorphosis as the game is played, changing the notion of 'in-frame' or 'out-frame' to the player. That dog barking in the distance might for example be the object of attention moments later in the game, as a consequence of the player's actions. However abstract as a concept, and difficult to implement well in the game design, the sound frame will always be present for the player, and audio game developers could do well with making it an object of interest.

Returning to the concept of the acousmatic situation, as discussed earlier, Beck (1998) shows that if it is put in relation to the sound frame, the notion of acousmatic sounds might differ between audio-only media and audio–visual media:

> "Acousmatic sound in radio drama is categorically different from film, because we hear every sound event in radio drama and connect it with its cause, within an overall sound space. Here, the acousmatic inhabits an area beyond the main frame of the sound picture, and it is 'out' and 'unseen', but of course is still bound by the sound space's 'outer frame'. I use the terms 'main frame' of the sound picture and 'outer frame' of the overall sound space."
>
> (Beck, para. 8.2)

On this, Forrester (2002) comments, "[n]otice the distinction made here between 'sound picture' and 'sound space', the first more object/image-like, the second event-determined and, needless to say, both in relation to an interpreting listener" (p. 45). Expanding on the ideas of Beck and Forrester, the principles of the acousmatic could therefore be related to the mental imagery process of the listener, where the listener identifies audible objects in the immediate surroundings and stores them within a mental picture—objectifying the sound events into a mental overview of the environment. Although the acousmatic is always related to hearing, the identification process could be done by any of the senses. Within most digital media productions, however, the concept of acousmatic sounds could be considered relative to the presence of visual content, and the conditions for when a sound is considered acousmatic changes depending on whether the visual content is present or not. In video games, the picture

frame already dominates the objectification process in the player's mental imagery. Acousmatic sounds in video games are therefore sound events that can only be objectified by the sense of hearing, such as sound events occurring outside the picture frame or many extradiegetic sounds.[21] In audio games, on the other hand, though being set in an all acousmatic situation, sound events inside the sound frame are objectified within the player's mental imagery, thus making them 'visible' to the player, and therefore, not truly acousmatic. Accordingly, it can be argued that only sound events representing game content that have currently *not* been objectified by the player can be described as acousmatic sounds in audio games.

As can be seen above, the player's imagination and ability to create a mental image of the virtual environment is of great importance in the gameplay experience of audio games. It is in the mind of the player that the game unfolds. Consequently, it may be shown that the limitation of the non-visual environment is actually one of the major narrative strengths of audio games. On a similar topic, Ferrington (1993) writes:

> "An effectively designed audio work may facilitate a listener's integration of life-based experiences into a 'movie' created within the 'theater of the mind'. Each individual becomes his or her own movie director with no two people having the same imaginary experience."
>
> (Ferrington, p. 62)

This mental process of visualization, referred to as *Sound Imagery* by some researchers (e.g. Forrester, 2002; Sonnenschein, 2001), is largely based on our individual cultural preconceptions and previous experience with sound events and their sources. Forrester (2002) writes that even though we perceive sounds as events, "many of the sounds we hear conjure up particular memories, associations and images of significant moments in the past" (p. 46).

As stated earlier, in a world of sound, it is important for the player to create a mental overview of the surroundings. Returning again to the causal listening mode, upon hearing a sound, the player tries to identify the cause of the sound. In this process, the player is also provided with information about the location of the source, as well as with additional information about the characteristics of the source—all of which will help the player's imagination in composing a mental image. A sound event could

[21] The concept of diegesis will be further explained in segment *3.3.3 – Diegesis in Audio Games*.

therefore be regarded as a sort of container, carrying important game related information to the listener. To exemplify, if an object undertakes an action and produces a sound, like the sound of a door closing, the player will be suggested a number of details, such as: (a) There is a door; (b) The spatial location of the door in relation to the player; (c) Characteristic features of the door (e.g. size, weight, material); (d) The fact that the door was open but now it is closed; (e) Acoustic information through reverberation of the surrounding environment (Is the door positioned in a large stone hall or in a small wooden room?); (f) Additional information about related events (e.g. Who closed the door? Why was the door closed? How was the door closed?); etc. The sonic landscape is a complex world filled with objects and on-going events, all of which will provide the player with similar types of information. Fortunately, the brain processes most of this information without much conscious effort on the part of the player.

Finally, sounds are events in time, with a beginning and an end. It would therefore be impossible to depict a sonic world with all its inherited objects at any particular moment in time. On this, Theo van Leeuwen (1999) wrote the following:

> "It is often said that vision allows, even prefers, things to be held still, so that they can be scrutinized in detail, dissected. […] Sound, by contrast, is seen as immaterial and evanescent: it is not, and can never be, an object which can be grasped and possessed. It can only be experienced and remembered."
>
> (van Leeuwen, p. 195)

This further emphasizes the need for the audio game environment to be held and remembered within the mental world of the player. Audio games can only depict the world at the immediate present. Since there is no freeze-frame picture of the environment to be found—only the current flow of aural events—the mental map becomes a necessity for the player.

3.3 Conditions of Sound within the Virtual Environment

The third entity of the audio game machinery is the *virtual environment.* When discussing films or other non-interactive media, it is common to talk about the *fictional* world, which then refers to the imagined world in which the characters and events of the narration is set.[22] However, Jørgensen (2009) emphasizes that the narrative world of digital games should not be confused with the term 'fictional world', and instead proposes to use the terms 'virtual world' or 'virtual environment' in relation to interactive media (p. 56). While fiction commonly refers to something imaginary, or something non-existing in the actual real world, the real world is still closely integrated with the imaginary world of digital games. With reference to Juul and Ryan's theories on imaginary and fictional spaces, Jørgensen argues that the boundaries between the possible imaginary worlds and the real world brake when the player's actual movements are integrated with the rules of the game system and the actions taken within the game world—as well as the consequences of these player–game world interactions—making the virtual reality of digital games a parallel potential reality for the player to pursue and realize (p. 56–59).

Since the virtual environment of audio games primarily consists of sound, this section will focus on the conditions required for sound to be utilized and perceived in an aural environment.

3.3.1 Sound as Events

Sound depicts events, not objects. This is a very important notion to comprehend when one aims to understand how sound functions in digital games. In explaining this concept further, Forrester (2002) states the following:

> "While vision is an object world, sound is an event world. Auditory experience is always a flow of sound, constant at times, rising or falling in intensity, noticeable when absent or excessive, but never truly 'silent'."
>
> (Forrester, p. 38)

[22] Oxford Dictionaries has two definitions of the word *Fiction*: (a) "*[mass noun]* literature in the form of prose, especially novels, that describes imaginary events and people"; (b) "something that is invented or untrue" ("Fiction," n.d.).

In the previous chapter, the consequences of this fact were discussed from a player's perspective, but this event-based world also has consequences from a game designer's perspective. When developing an audio game, much care and thought have to be put in when designing the soundscape in order to accurately describe the game world for the player. The environment is full of objects, but as stated earlier, sound cannot depict objects, only events in relation to objects. For example, a car does not itself make a sound, but a car running its motor does. Forrester (2002) clarifies this further by using wind as an example:

> "To hear a sound is to hear the cause of the sound. The wind is a good example particularly as it is the effect of 'the wind' on objects which constitutes the sound of the wind as an event or events; for example, compare the sound of leaves blowing with the intermittent but continual sound of a door swinging and banging – both caused by the wind."
>
> (Forrester, p. 37)

Consequently, when depicting the world of objects, audio game designers need to fill the virtual environment with events. Related to this is an interesting reflection: when one talks about an object being *invisible*, most people will likely think of it as having magically vanished from perception. If, on the other hand, an object is *inaudible*, it most often means that the sound energy emitted by the object is too weak to be heard but that the object is still very much there, perceptible for most people—only not by hearing.[23] In an auditory world, however, a silent object is an invisible object—hidden, secret, unperceivable. All objects in the virtual environment must generate a sound if it is to be perceived by the player. However, if the virtual soundscape becomes too crowded, the sounds start to mask each other and become difficult to discern. In his seminal work on acoustic ecology, Schafer (1977/1994) coins the terms *Hi-fi* and *Lo-fi* soundscapes, to be used while describing sonic qualities in natural acoustic environments:

[23] By comparing the inputs on www.collinsdictionary.com we find that 'invisible' is synonymous with words such as *unperceivable, hidden, secret, concealed, unnoticeable* etc, whereas 'inaudible' is synonymous with *indistinct, low, stifled, mumbling, unheard* and *out of earshot.* ("Invisible," n.d.; "Inaudible," n.d.)

> "A hi-fi system is one possessing a favourable signal-to-noise ratio. The hi-fi soundscape is one in which discrete sounds can be heard clearly because of the low ambient noise level. The country is generally more hi-fi than the city; night more than day; ancient times more than modern. In the hi-fi soundscape, sounds overlap less frequently; there is perspective—foreground and background […] In a lo-fi soundscape individual acoustic signals are obscured in an overdense population of sounds. The pellucid sound—a footstep in the snow, a church bell across the valley or an animal scurrying in the brush—is masked by broad-band noise. Perspective is lost."
>
> (Schafer, p. 43)

When a developer designs an audio game environment, the amount of information present will always have to be balanced against the comprehensibility of the soundscape. These factors, however, are not linearly related. While a lo-fi soundscape could be regarded as an information overload, due to its multitude of sounds, the information discernible by the player might actually be radically less than in a more hi-fi soundscape. The hi-fi soundscape, while presenting fewer sources of information, might still present more information *detail* to the player than the lo-fi one. In the end, the fidelity level depends more on the quality of the sound design rather than the amount of sound sources used.

3.3.2 Balancing the Sounds

In his handbook on mixing music, Izhaki (2010) divides the mixing process into the balancing of sounds according to the domains of time, frequency, level and space (p. 59). The *time* domain refers to the temporal qualities of a sound, such as duration or rhythmic structure. A sound event always has a beginning and an end, after that there is only silence—and, as mentioned earlier, a silent object is not perceivable in the auditory world. Therefore, audio game designers have to carefully plan the temporal aspect of the soundscape, making sure a sound event is heard as long as the source is present, while still making time for other events to be played. Balancing sounds in the *frequency* domain means placing and processing the sounds according to their frequency content, in order to present a fuller sounding soundscape and achieve intelligibility by reducing masking effects among sounds. When mixing sounds according to the *level* domain, the engineer balances the energy volume (amplitude) of

involved sounds, making sure all sounds are audible while perhaps presenting certain sounds as foci of interest. The *space* domain is, according to Izhaki's definition in music mixing, formed of the two subdomains *stereo* and *depth* (p. 59), which are used to describe a sound's placement and width in the perceived soundstage of a stereo mix. Following the same general principal, the *space* domain in three-dimensional games would relate to the spatial placement of sound events within the virtual environment, and in some cases, also the perceived size of the source. To avoid causing an imbalanced fatigue on the ears of the player,[24] sound events should—preferably when possible—be placed somewhat evenly in the spatial plane of the virtual environment, in regards to their other attributes of time, frequency and level.

When mixing a sound design for games, the above-mentioned domains are all present, but the designer also has to take into account the interactivity aspect introduced by the presence of a player, and the dynamic nature of a game system with randomized events. In relation to this ever-changing system of cues and surprises in digital games, Grimshaw (2007) writes about the implementation of *aural cues*:

> "These are contrived and calibrated by the sound designers and game designers in an attempt to produce the desired mix of both predictable and unpredictable outcomes of player responses that provide the patterned, yet differing, gameplay experiences each time the game is played."
>
> (Grimshaw, p. 121)

Ferrington (1993) writes that the "goal of good audio design is to effectively engage the listener in active and attentive listening" (p. 67). In relation to listening, Schafer (1977/1994) writes that in the "ultimate lo-fi soundscape the signal-to-noise ratio is one-to-one and it is no longer possible to know what, if anything, is to be listened to" (p. 71). On the other hand, in a hi-fi soundscape, vital or interesting information can be communicated by even the slightest disturbance (p. 43). Considering Schafer's general ideas that the hi-fi soundscape more easily enables the listener to engage in active and attentive listening than the lo-fi one, a conclusion could be drawn that the listener's attentiveness relates directly to the fidelity of the soundscape. In a hi-fi environment, the listener hears less but listens more, whereas in a lo-fi environment, the listener hears

[24] A purely theoretical scenario could, for example, be walking for a few minutes down a long corridor with a mosquito buzzing around ones right ear and an elephant on a leash walking on ones left side.

more but listens less. In audio games, where sound presents the only source of information, the quality of that information becomes especially important in engaging the player's attentiveness. Combining Ferrington's and Schafer's theories, a hi-fi soundscape could then be said to be preferable in the virtual environment of audio games.

If we continue to draw a parallel between soundscape fidelity and the attentiveness of the listener, the relationship between the lo-fi soundscape and passive hearing needs to be explored further. To do this, we need to return to the discussion on listening modes presented in the previous section, where, with reference to Jørgensen (2009), this thesis utilized the listening modes of Chion to describe active listening processes, and the views of Smalley to explain the passive hearing process. Smalley (1996) bases his theories on aural perception on the relationship between the subject who is perceiving (the listener) and the objects of perception (the sounds) (p. 81–82). In his description of the reflexive relationship, Smalley writes the following:

> "The reflexive relationship is taken from Schachtel's autocentric mode. It is subject-centred and is concerned with basic emotional responses to the object of perception. The object has no real identity separate from the subject's emotion. This type of relationship can be active or passive but has a strong tendency towards passivity since there is little or no exploration of the object of perception, only response to it."
>
> (Smalley, 1996, p. 82)

Accordingly, when a person passively hears a sound event, no active exploration of the sound properties will occur, but the person will still respond to it on a subconscious, emotional level. Considering the previous claim that the lo-fi soundscape induces passive hearing on the listener, an audio game developer could perhaps utilize the reflexive relationship as a narrative tool. If designed well, intentionally implementing a more lo-fi soundscape in specific areas of the game might stimulate certain emotional effects on the player—though, possibly at the cost of conscious intelligibility.

On a last note on this topic, I would warmly welcome more research in this area of soundscape fidelity and listener attentiveness, especially to conceptualize listening modes adapted to the playing of digital games. Further studies of the theories of listening presented by researchers such as Chion, Smalley and Schaeffer, in

combination with the notion of soundscape fidelity and an interactive audio environment, will surely bring more extensive knowledge to the research field of game audio as a whole.

3.3.3 Diegesis in Audio Games

Present in all narrative media productions is the concept of *diegesis*. The diegetic notion is an analytical tool that explains the various narrative channels and dimensions present in a fictional production. Quite simply, diegesis can be used to define and separate the origins and aims of information, and categorize these correspondingly within various planes of the narrative. In film theory, the most commonly used levels of diegesis are the *diegetic* and the *extradiegetic* levels. The first level refers to the dimension internal to the fictional universe, while the second level refers to the dimension external to the fictional universe.[25] A third, less commonly known level, is the *hypodiegetic* level, which is sometimes used in literature studies to refer to a 'story-within-a-story' (Malina 2002, p. 145). Jørgensen (2009) further explains the concept of diegesis in relation to sound:

> "*Diegetic sound* originates from within the fictional universe, where we assume that the characters perceive the sound as natural to that world, while *extradiegetic sound* is a commenting aesthetic feature with no direct connection to an actual source within the fictional universe, and consequently, the characters do not seem to be able to hear it."
>
> (Jørgensen, p. 98)

Returning to the notion of the player–avatar relationship, as discussed in the previous section, it was stated that the function of that connection is to convey the sensual impressions of the avatar to the player. Although the player and the avatar belong to separate entities in the game machine, this connection transcends information between the two, merging them together as two very differentiating sides of the same coin. In exploring this connection, Jørgensen (2009) identifies the concept of *transdiegesis.* In film theory, the diegetic and the extradiegetic levels typically function independently from each other. However, as Jørgensen points out, the presence of the player makes it

[25] The *extradiegetic* level is also commonly known as the *non-diegetic* level. This thesis, however, uses *extradiegetic*, as the term *non-diegetic* implies 'anything other than diegetic', which is confusing when further levels of diegesis are added.

impossible to directly apply film theory's notions of diegetic and extradiegetic, without also identifying that the one can influence the other (p. 99). Jørgensen explains this notion more explicitly, stating the following:

> "[The] double player position allows computer games to utilize extradiegetic sound to provide the player with information relevant for game-internal choices of actions. This leads to the interesting situation that although the game character does not hear extradiegetic sound due to its game-internal position, in effect it may react to extradiegetic sound because of the direct control link between player and avatar. In this respect, the game character can evaluate and act upon information that it should not be able to hear."
>
> (Jørgensen, p. 99)

Accordingly, the concept of transdiegesis is not a diegetic level itself, but rather refers to a situation that commonly occurs in digital games when the functional information of a game object transcends the border of its own level of diegesis, to directly or indirectly affect the events and actions on another level of diegesis. Jørgensen's (2009) theories on the principles of diegesis in video games apply seamlessly to audio games as well. In games without a visual interface, however, the line between of what is meant for the avatar or for the player might sometimes be difficult to distinguish. In the audio game Papa Sangre (2011), for example, a voice guides the player through the game with semantic information about the game plot, gameplay challenges and descriptions about the environment. In the first level of the game, this voice presents itself as 'a watery fluttery thing', and based on the textual content, it seems as though this character is somehow present within the diegetic game environment.[26] However, the voice on many occasions speaks about information only known to the player, such as control-related information. This character can therefore be seen as placed within a transdiegetic plane, with one 'foot' in the game world and the other in the extradiegetic dimension.

Judging from the speech by 'a fluttery watery thing', it seems, however, that the developers consider the player and the avatar to be one unit in Papa Sangre (2011).[27]

[26] See transcript in Appendix A.

[27] See transcript in Appendix A.

It is the player that has ended up in 'the kingdom of Papa Sangre', and the 'new body' of the player–avatar unit is represented by the game system (the iPhone), offering the game an ultra-immersive quality.

3.3.4 Final Reflections on the Characteristics of Sound in Audio Games

In video games, sound is a very important, albeit not vital entity. The visual expression alone often generates enough stimuli to enable the player to participate in the virtual environment. However, as Sullivan & Gil write, "sight paints a picture of life, but sound, touch, taste and smell are actually life itself" (as cited in Rodaway, 1994, p. 4). In Jørgensen's (2009) study, sound is described as contributing a sense of *presence* to the game world:

> "Sound is described as a surrounding feature that tells the player about the spatial layout of the environment and what goes on in it. To feel present in a space requires the possibility to sense the space around oneself, and this feeling is ensured by enabling the player to hear sounds also from somewhat distant sources."
>
> (Jørgensen, p. 204)

This is also supported by the previous claim that sound only depicts events, not objects. Objects *can* be the source of a sound, but a dead, motionless object will never generate a sound. On a similar topic, Schafer (1977/1994) writes that the sense of touch is the most personal and intimate of our senses, and since both touch and hearing perceive mechanical vibrations, Schafer views the sense of hearing as a way of "touching at a distance" and argues that "the intimacy of the first sense is fused with sociability whenever people gather together to hear something special" (p. 11). In audio games, where sound is the only source of sensual representation, it also has to compensate for the lack of other senses. As Schafer argues, hearing and touch are closely related, and the two senses "meet where the lower frequencies of audible sound pass over to tactile vibrations (at about 20 hertz)" (p. 11). Accordingly, the lower the frequency content of a sound event is, the closer it relates to the sense of touch. In digital games, many sound events are intentionally enhanced in order to create an emphasis of physical contact with the avatar, such as a sound event describing a powerful punch being dealt to the player. Even though a realistic reproduction of such an event would sound fairly weak, by augmenting the lower frequencies, the powerful punch can almost be 'felt' by the

player's sense of hearing. Other senses may be more difficult to compensate in a similar manner, but as has been shown previously in the thesis, the player's mental imagery ability can be a powerful tool in compensating for lack of visual stimuli. And finally, through semantic use of verbal language, sound is also able to produce arbitrary descriptions of almost anything, including the experience of a certain smell or taste perceived by the avatar.

On a final note, Schafer (1977/1994) writes that for any sound to be heard, it must first overcome the inertia of the eardrum to start vibrating, causing minor imperfections and distortions to the perceived sound (p. 262). Only a sound initiated before our birth and continued until after our death would be free from such distortion—a *perfect* sound. However, as Schafer illustrates, such a sound, "initiated before our birth, continued unabated and unchanging throughout our lifetime and extended beyond our death, would be perceived by us as—*silence*" (p. 262). According to this, Schafer concludes the following:

> "[A]ll research into sound must conclude with silence—not the silence of negative vacuum, but the positive silence of perfection and fulfilment. Thus, just as man strives for perfection, all sound aspires to the condition of silence."
>
> (Schafer, p. 262)

One of the finer qualities of sound is that it is omnipresent. It is ever present in space, constantly flowing in time, and surrounds anyone who can hear it. Only in a virtual environment would non-omnipresent representation of sounds be possible, as virtual environments do not share real world environment laws, only simulates them. As was shown in this chapter, the avatar acts as an extension of the player's senses within the game world environment. The avatar itself, however, does not have real ears, or eyes, or any other physical function of reproducing sensual stimuli. Yet, it is through the avatar's perspective that the diegetic world is presented. The avatar somehow 'magically' perceives its environment, and only in the transducing process of the game system is the signal distorted. Reflecting on Schafer's thoughts, it can perhaps be concluded that the avatar hears the sound of its environment as perfective—or maybe, the only truly perfective aural world is the virtual environment with no sound events in it.

4 The Various Functions of Sound in Audio Games

Now, as the context within which sound is placed in audio games has been explored, it is time to focus on the various *functions* of sound. In order to define *what* the role of sound is within audio games, an explanation of *how* it acts and relates to the game machinery is needed. To make that explanation, this chapter will draw knowledge from two main areas of research: (a) Michael Halliday's approach to communication according to the field of systemic functional linguistics; and (b) Kristine Jørgensen and Sander Huiberts & Richard van Tol's different theories on game audio functionality. In applying these theories to the field of audio games, this research aims to gain knowledge of how sound functions in audio games, and in which ways this might differ from the functional uses of sound in video games.

The first section of this chapter will discuss the concept *metafunctions of communication,* as proposed by Halliday (1978). The second section will present theories on the specific functionalities of sound in computer games, as presented by Jørgensen (2009) and Huiberts & van Tol (2008, 2010). Finally, the last two sections of this chapter will present some of the more audio game specific usages of sound and technology that were identified during this research, and explain these in relation to the theories from the first two sections.

4.1 Metafunctions of Communication

As stated in the previous chapter, sound events are containers of a large variety of information. Within digital games, this information—originating from any of the diegetic levels—communicates within its native dimension, and in some cases, also transcends the diegetic borders, establishing communications with other levels of the diegesis as well. Regardless of on what or between which levels of the diegesis this communication takes place, the information ultimately reaches the ears of the player, who in turn acts both as an observer and an active communicator with the game. As can be seen, to map out an overview of the communicative channels in digital games will quickly require some abstract thinking.

As a system to use for defining the functional components behind language, Halliday (1978) proposes three *metafunctions* that are simultaneously present in all communication: the *ideational* metafunction, the *interpersonal* metafunction and the *textual* metafunction (p. 112). Although Halliday identifies these within the use of

language, the metafunctions can be used to analyze any form of social communication (Van Leeuwen, 1999; Wingstedt, 2008), such as digital games—where communication of information takes place on multiple levels simultaneously. In all forms of communication, meaning is created by both the communicator and the interpreter (receiver). In the context of verbal language, the speaker is often considered the communicator. In the complex social system of digital games, however, the communicator could instead be any part of the game machinery, such as the player, the game system itself or any object or event within the virtual environment. However, Barbosa, de Souza & de Paula (2003) argue that according to semiotic engineering, digital games represent the "designers' response to what they believe are the users' problems, needs, preferences and opportunities" (p. 18). They continue to describe the games as a communicative extension of the designer by stating the following:

> "In fact, user–system interaction is viewed as a conversation between the user and the "designer's deputy" – a *programmed* representative of the designer. Since the designers are no longer there at the time such conversations are carried out, they must anticipate every possible conversation. Thus, the designer's deputy is capable of participating in all and only the designed conversations."
>
> (Barbosa, de Souza & de Paula, 2003, p. 18)

Accordingly, when a game is playing, scenarios of communication will still take place, but since these scenarios have at one point already been designed, the game designer could ultimately be regarded as the communicator in digital games.

In language, the communicator is presented by an inertia to be overcome, present in the obstacle of the language itself. No communication is perfect; there always exists an alteration of meaning between the participants of the communication. In addition, meaning is also formed by the interpreter's own interests—as the interpreter often 'hears what he or she wants to hear'. In relation to this, Halliday (1978) coins the concept *meaning potential*, which refers to the potential to observe, convey or form variations of meaning by the participants in the communicative act.

In any social context, there will always be a discrepancy between the meaning created by the communicator and the information presented by the environment (or by other communicators). Incorporating this, the *ideational* metafunction relates to the communicator's potential of making meaning to the information observed. According

to Halliday (1978), the ideational metafunction "is the component through which the language encodes the cultural experience, and the speaker encodes his own individual experience as a member of a culture" (p. 112). The ideational metafunction can therefore be regarded as the part of communication in which information about the communicator's environment is processed and shaped according to the communicator's own experience or properties. In digital games, this can relate to any information in or about the virtual environment; such as game objects, events, status indicators, ambient sounds, etc.

Relating this to the discussion on sound imagery in the previous chapter, the ideational metafunction can pose very different experiences between communicators and interpreters of audio games. Because of the lack of visual stimuli, the mental process in which the player objectifies game objects and creates an overview of the game environment is to a larger extent based on previous personal experiences in audio games than in video games. When the game is being played, the communicator (the game designer) has already encoded the game world through the available tools according to his or her own subjective perception—in turn leaving the players to interpret the same game world according to *their* subjective perception of it. In an interview on YouTube, Nick Ryan, director of sound and music of the audio game Papa Sangre (2011), further explains this issue by saying that "there is a huge amount of subjective interpretation among our players, so it's quite important for us to be sure that what we as game designers think we are making is what the players perceive it to be" (BAFTA Guru, 2012, 1:54).

The *interpersonal* metafunction relates to the communicator's potential in communicative interaction with others. Halliday (1978) writes that the interpersonal is "the component through which the speaker intrudes himself into the context of situation, both expressing his own attitudes and judgements and seeking to influence the attitudes and behaviour of others" (p. 112). The interpersonal metafunction concerns relationships among the participators of communication—how these relationships are formed and how they relate to each other.

Since digital games are by nature interactive, there will always be an exchange of meaning between the player and the game. The interpersonal metafunction can be used as an analytical notion to describe these relationships between the game system, the virtual environment and the player. For example, when a sound engages the player by proposing a question or a statement that demands a reaction from the player (such as

the sound of a threatening object), the interpersonal relationship shows that the game may have a dominating status over the player, effectively forcing the player into taking action. Van Leeuwen (1999) writes that "[s]ound does have its ideational resources, but they have to be realized on the back of interpersonal resources" (p. 191). The interpersonal metafunction is therefore also in effect when a sound induces a semantic or affective meaning based on the player's contextual experience (such as affective music or voice messages), since a relationship exists between the sound and the player.

The *textual* metafunction refers to the communicator's potential in forming a coherent structure through an expressional mode, such as writing. Halliday (1978) writes that the textual metafunction "is the component which provides the texture; that which makes the difference between language that is suspended *in vacuo* and language that is operational in a context of situation" (pp. 112–113). As Halliday's usage was intended for language, this would for example relate to the way we form words into sentences, sentences into paragraphs and paragraphs into sections. The overarching structure of communication is just as important as the specific content. For example, a situation containing 'a dog', 'a rabbit' and 'dinner' may be structured in various ways. The sentence 'the dog ate a rabbit' has a completely different meaning if the structure of the words is changed into 'the rabbit ate a dog'.

In digital games, the stream of sound provides a textural continuity and form to the time dimension. When creating a scene, or a level in the game, sounds are used to organize time and space, in order for the player to understand the scene as a whole. For example, when the player walks from one location in the game to another, the soundscape will change accordingly, making the player aware of the change of scenery. In addition, music can be used to overlap this change of scenery. For example, if an enemy chases the avatar, the avatar may move from one location to another, while extradiegetic music continues to present the sense of danger to the player until the chase is actually finished.

4.2 Functions of Sound in Digital Games

The two most comprehensive research theories of game audio functionality found during this study were formed by Kristine Jørgensen and Sander Huiberts & Richard van Tol. Although having slightly differentiating approaches and results, their theories support each other by presenting two different analytical tools of game audio—equally useful depending on the angle of research. This section will briefly present both theories, in order to provide a more comprehensive background to the functions of game audio.

4.2.1 The IEZA-model by Huiberts & van Tol

In his study on sound as an immersive design tool in video game development, Huiberts (2010) presents a graphical model of sound event analysis named *IEZA.*[28] Within the IEZA-framework, Huiberts identifies four domains in which to categorize sound events in games, based on their communicative functions and diegetic placement: *Interface*, *Effect*, *Zone* and *Affect* (pp. 24–29). These domains are described in more detail in the original thesis, but will here be summarized (based on descriptions by Huiberts, 2010):

- Interface-sounds originate within the extradiegetic level, with the function of communicating usability information about the gameplay, such as menu item feedback or information about health status and score points (p. 28).
- Effect-sounds originate within the diegetic level, with the function of communicating information concerning objects and events present in the game world, such as footsteps, dialogues, weapon sounds or other sounds triggered by the avatar or any other diegetic object or event (pp. 25–26).
- Zone-sounds originate within the diegetic level, with the function of providing an ambient background layer of sound, communicating scenic information about the game environment, such as weather sounds, crowd-murmur or environmental noise (pp. 26–27).
- Affect-sounds originate within the extradiegetic level, with the function of communicating affective information, such as music or extradiegetic ambient-like FX sounds (often used in horror games) (pp. 28–29).

[28] The IEZA-framework was originally developed by both Sander Huiberts and Richard van Tol and was published in 2008 on www.gamasutra.com. For clearer references, however, this thesis will hereafter refer to Huibert's doctoral thesis 'Captivating Sounds' (2010).

The relationships between these four domains are further clarified in the graphical representation of the IEZA-model, as illustrated below in Figure 1:

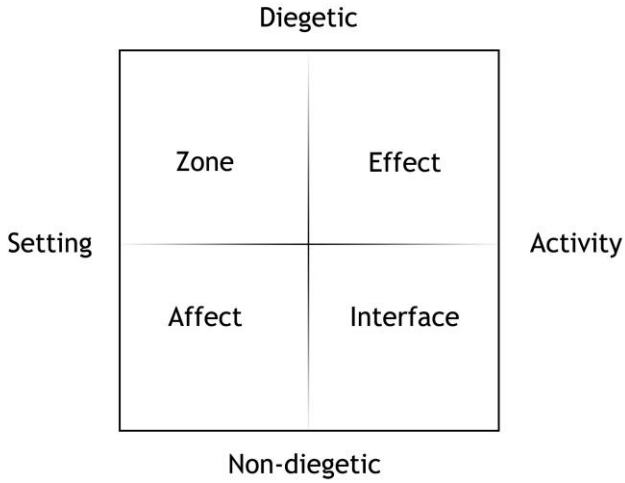

Figure 1. The IEZA-framework, used for game audio categorization.[29]

In addition to the four domains, Huiberts (2010) also includes two intersecting dimensions in the IEZA-framework: *diegesis* and *interdependence*. In the dimension of diegesis, Huiberts (see Figure 1) links the 'Zone' and 'Effect' domains together in the diegetic level and the 'Interface' and 'Affect' domains together in the extradiegetic level (pp. 22–25). Simultaneously, in the interdependence dimension, Huiberts relates the domains ''Interface' and 'Effect' to provide information about the *activity* of the gameplay, whereas 'Zone' and 'Affect' provide information about the *setting* or mood of the gameplay (p. 24).

By identifying the relationships between sound events with the IEZA-framework, Huiberts (2010) suggests two overarching themes of main usages of sound in video games:[30] (a) sound can be used to *optimise* the gameplay experience by providing important information that assists the player in playing the game; and (b) sound can be used to *dynamise* the gameplay experience by stimulating the player on an emotional level with enriching and thrilling sounds, such as music or ambient sounds (p. 29). Although the optimising usage relates more to activity sounds, and the dynamising usage relates more closely to setting sounds, Huiberts emphasizes that this is not always the case, as sound events often carry many functions at the same time (pp. 30–31).

[29] Used with permission from Huiberts & van Tol.

[30] Huiberts defines these main usages as the two *Roles* of sound in games. However, as a matter of definition, I will wait to use this term until section *5.3 – Analysis of Sound as a Function and a Role*.

4.2.2 Jørgensen's Four Functions of Sound

In her research, Jørgensen (2009) identifies the main functional uses of sound in computer games and divides these into four categories: *Usability*, *Atmosphere*, *Orientation* and *Identification* (pp. 158, 244–246). Sound events in the *usability* category are sounds that directly relate to the player as an active participator of the game narrative. Jørgensen describes these as "sounds that provides the player with information relevant for actions and events in the game" (p. 244). Further on, she states that these sounds are either *reactive*, when they appear "immediately after a player action and works as a response signal", or *proactive*, when they work "to provide urgent information about an upcoming situation that the player needs to evaluate" (p. 244). For example, a gunshot sound may be *proactive* if the gun is fired by an enemy and the player is now urged to take cover, or *reactive* if the player fired the gun and the sound is a positive confirmation that the gun was indeed fired. A reactive response may also present a negative rejection, as would be the case if the player upon firing the gun instead hears a clicking sound—symbolising that the gun is out of ammunition and that the gun did not fire. In the *orientation* category, Jørgensen catalogues sounds that contain spatial information about objects, events and situations within the game environment (p. 158). These are sounds that help to orientate the player in the game world by providing information on where and how certain events, objects or situations occur. (p. 245). Jørgensen writes that sounds with an *atmospheric* function work to add certain moods to the gameplay experience, in order to increase the player's engagement level and create a sense of life and presence in the game universe (p. 158). Jørgensen mentions that while usability sounds often also carry an atmospheric function, many atmospheric sounds are not connected to usability, such as ambient sounds or linear music (p. 245). Lastly, Jørgensen states that sounds within the *identification* functional category enable the player to identify specific information about the source itself, and compare this to information gained from other sources—in order to evaluate their relative values and make an active decision on how to proceed in the game (pp. 158–159). Although some of the above-mentioned functions can relate to the various modes of listening or hearing—as discussed in the previous chapter—the *identification* function might be the one that most closely relates to Chion's notion of *causal listening.* It should be noted though, that Chion's listening modes define perceptual approaches from a listener's perspective, whereas Jørgensen's functions define the meaning and activity-related purpose of a sound in relation to the player.

Additionally, Jørgensen (2009) writes that these functions often overlaps and work together in a synergetic manner (p. 159). For example, a sound event representing an aggressive roar would have a clear proactive *usability* function, in terms of urging the player to take action ('fight or flight?'). Simultaneously, it would possess an *atmospheric* function, by providing a sense of presence to the game world and adding a fearful mood to the situation. In addition, it would also provide the player with *identifying* information about the source (e.g. 'is it a large monster or a small dog?'), as well as *orientational* information about the source's spatial position in the environment.

4.3 Virtual Environment Features in Audio Games

This section will analyze some interesting usages of sound in audio games that are, if not unique to the audio game form, at least uncommonly used within video games. These examples, however, are just a small list, based on my own observations on some of the games studied during this research. Since the audio game scene has so far been rather experimental in its game development, many more unusual and creative usages of sound can surely be found—or implemented in future developed games.

4.3.1 Assistive Navigational Features

Of all the unique usages of sound identified during this research, it seems that most of them in one way or another relate to assisting the player in navigating through the virtual environment. Since most people rely primarily on vision to navigate in their environment, navigation in an audio-only environment might show to be a difficult task. Even players that are experienced in navigating through audio-only environments may often need extra information about the game world in addition to the information provided by the presence of common diegetic game objects—just as video games often includes assistive visual features, such as maps that present an overview of the environment or various graphical pointers that show the player where to go. Many audio games have therefore implemented certain aural features, which help the player understand how the environment is set up and make the gameplay more accessible and rewarding. These features will in this thesis be referred to as *Assistive navigational features*.

In the audio game Swamp (2012), an assistive navigational feature is found in a game object simulating a sonar-device, which provides spatial information about objects in the environment, such as walls. When pressing a certain keyboard button, a

'beeping' sound is perceived as being 'sent off' in all directions, and if an object is present in its path, the sound gets louder, simulating that it is reflected back upon the avatar. It is not clear whether this sound belongs to the extradiegetic aural interface of the game or if it has an actual reference within the game world (i.e., if the avatar is in possession of this sonar-device). It does, however, simulate acoustic sound projection and reflection within the environment, and provides the player with information about diegetic objects. Accordingly, this sound could be categorized as an *effect* sound (Huiberts). Since the sonar sound provides the player with spatial information about objects in the game world, it could be said to have an *orientation* function (Jørgensen). However, in digital games, most sound events with an orientation function state the presence of the sound's source within the game world, along with other functional information about the source. In those cases, the orientation function is subsequently a result of the presence of the avatar and its relative position to the source of the sound. The sonar sound in Swamp, on the other hand, is an effect sound that provides the player with spatial information about *other* objects. It does not present any other information about the objects except for stating the presence of the objects and their relative position to the player. Furthermore, the sonar sound can also be said to have a *usability* function. Not only is it *reactive* on the player's key command, but it also activates an indirect *proactive* function in the objects reflected by the sound. By marking the presence of a wall, this sonar sound can advise the player from walking further in that direction, whereas denoting a hole in the wall could urge the player to go through it. In video games, only denoting the presence of an object would not necessarily ascribe usability functional properties to a sound event. However, when the player is dependent on sounds to navigate, orientation sounds can provide the player with valuable information related to usability.

Another assistive navigational feature of sound relates to the avatar's location and facing direction within the game environment. Also in the game Swamp (2012), the player is provided with the feature of an aural compass. Upon holding down a certain keyboard button, a voice is heard circling around the head of the player (through headphones), denoting the cardinal points as it goes. The direction that the avatar currently faces is then expressed somewhat louder than the others. To exemplify, if the avatar is facing east, the voice would say "EAST – south – west – north – EAST – south …" and so on until the button is released. In this case, 'east' would be perceived loud in the centre of the stereo image, whereas 'south' would be slightly lower in volume,

coming from the right stereo channel; 'west' in the centre, also in low volume (indicating to originate from behind the player); and 'north' on the player's left side, also in low volume. Like the sonar sound, this feature can be categorized as either a diegetic *effect* sound or an extradiegetic *interface* sound (Huiberts), depending on the contextual explanation. Furthermore, it has a clear *orientation* function (Jørgensen), describing the spatial *interpersonal* relationship between the avatar's position and the game environment (Halliday). Moreover, it also resembles the sonar sound by having a *usability* function (Jørgensen), with both *proactive* and *reactive* properties. Another example of a similar feature is briefly described in a research article by Friberg & Gärdenfors (2004), in which the authors state that they—the developers of the audio game Tim's Journey—have implemented the sound of a foghorn in every quarter point of the compass in their game (p. 150). However, since this game does not appear to have been published, no more detail about the specific functional properties of this feature has been found.

Yet another usage of sound related to navigation, is the interiorizing feature in the game BlindSide (2012),[31] where the avatar describes objects in his surroundings to the player. In this feature, the developers use semantic meaning to describe information about game objects and the environment to the player. Essentially, it functions in the same way as the sonar: the sound event provides information about objects and events, even though the source of the sound (the avatar) is not directly linked to the object of focus. However, in addition to providing orientation and usability information about the objects—like the sonar sound does—this feature also provides identifying and even atmospheric information about the avatar's surroundings.

The semantic use of voice as an assistive guiding character is a common feature in many audio games. In Papa Sangre (2011), there is the 'watery fluttery thing';[32] in BlindSide (2012), there are two guiding characters—an unknown male voice that mainly provides extradiegetic information, and the avatar, which provides diegetic information; in Swamp (2012), various levels of status information can be read out loud when pushing certain keyboard buttons—such as information about the player's health, the amount of ammunition remaining in the current weapon or the current location of the avatar; etc.

[31] Described earlier, in segment *3.2.1 – The Player–Avatar Relationship*.

[32] Described earlier, in segment *3.3.3 – Diegesis in Audio Games*.

Although it is not a sound event itself, Friberg & Gärdenfors (2004) describe another assistive navigational feature within their audio game Tim's Journey, named the *Ambience Reductor*, which allows the player to "temporarily reduce the volume of all sound that he/she cannot directly interact with" (p. 150). This could be a very useful feature in many audio games. Although it would affect the *dynamising* function of sound in a negative way, the ambience reductor could greatly improve the *optionalising* function (Huiberts), helping the player to progress in the game if stuck.

4.3.2 Direct and Indirect Functional Usages

As can be seen in the exemplified features above, Jørgensen's (2009) theories on game sound functionalities are generally applicable also to these, somewhat rare, game audio implementations. However, Jørgensen's research mostly studies direct usages of the functions, meaning that the analyzed sound event presents functional information about the source of the sound event itself. The exemplified features above, on the other hand, demonstrate that it is also possible for a sound event to present functional information about *other* game objects, not directly related to the source of the analyzed sound event. In relation to Jørgensen's functions, I therefore suggest the addition of the factors *direct* or *indirect* usage, which can be applied to any of these functions. To clarify this classification further, an exemplified usage of sound is presented within each functional category below in Table 1:

	Direct usage	*Indirect usage*
Usability function	A weapon sound that confirms or rejects the firing of a gun.	A sonar sound that confirms or rejects the presence of a wall.
Atmosphere function	A wind sound that provides an ambient layer to the game.	A semantic use of voice that provides atmospheric information about the environment (e.g. "Oh, it's cold in here!").
Orientation function	A motor sound denoting the position of a car in relation to the avatar.	A compass sound denoting the avatar's facing direction in relation to the cardinal points.
Identifying function	A barking sound describing a large, fearful dog.	A semantic use of voice describing objects in the environment (e.g. "That dog sure is big!").

Table 1. Examples of *direct* and *indirect* usages of Jørgensen's functions.

By adding these two factors, the assistive navigational features can be distinguished from ordinary diegetic sound events by their indirect usage of Jørgensen's functions. The features presented in this section also show that semantic use of voice can be an efficient way to provide indirect functional information about the game environment.

To summarize, the *assistive navigational features* are not functions themselves, but rather certain usages of sound in digital games that help the player navigate through the unseen landscape of audio games. Although analyzable by Jørgensen's functions, they are characterized by providing functional information about *other* objects and events, in addition to the information about the analyzed sound event's own source. Whether a sound event provides direct information about its own source or indirect information about other objects or events in the environment can be an important distinction to consider when the functional properties of a sound event are being analyzed. Since these factors are only briefly explored in this thesis, further research on this theory is warmly welcomed.

4.4 Game System Features in Audio Games

In addition to some unusual usages of sound, there are also a number of creative usages of technology present in many audio games. Although these features are not exclusive to the audio game form, they may present different effects in audio games than they would in video games. This section will briefly list some of the rare technological features found in audio games during this research.

4.4.1 Gyroscope

In the game BlindSide (2012), a gyroscope sensor in the game system (mobile phone) provides the player with the option of rotating the avatar's facing direction by rotating the phone. Although this is not a unique technological usage for audio games, it may provide a deeper sense of ultra-immersion in audio games than it would in video games. If the same feature was implemented in a F-P 3-D video game, where the visual image on the screen of the phone only comprises a minor part of the whole perceptual field, the graphical environment would clash with the player's simultaneous perception of real world elements, dislodging the virtual environment from the player's immediate present and thereby interfering with the illusion of ultra-immersion. In the audio game BlindSide, however, the player only needs to sit in a rotatable chair, close the eyes, and *feel* the aural virtual world moving around him or her.

4.4.2 Head-Tracking Headphones

A technological feature that is similar to the gyroscope is found in a research paper by Röber & Masuch (2005), where the authors are experimenting with so called *head-tracking* headphones in audio games. The general function of head-tracking audio systems is to send information about the movements of the head from a motion sensor to a receiver, which then processes the audio signal to rotate the sound field in compensation of the movements of the head.[33] As perceived by the listener, when moving the head sideways with a head-tracking audio system, the sound field will stay fixed in relation to the listener's external environment. When listening to headphones without a head-tracking system, the sound field will instead appear to be originating from inside the listener's head, as the internal sound field will not change even when the listener moves the head sideways. Although this technology is applicable for video games as well, the position of the screen will stay fixed in the external environment, causing the player to only make slight movements of the head in order not to loose sight of the screen. In addition, if the player uses head-tracking headphones while playing video games, conflicting stimuli inputs may occur, as the aural sense representation is rotatable separately from the fixed visual perspective of the avatar on screen. In audio games, on the other hand, the virtual environment is only represented through aural means, and therefore no stimuli conflicts will occur when rotating the head using head-tracking systems. It should be noted though, that these head-tracking systems are bought entirely separately from the games themselves, and are still rather uncommon among players of digital games. The features of this technology could, however, be very useful in the experience of an audio game, or as Röber & Masuch (2005) conclude, "[t]he use of head-tracking allows a much more natural interaction with the acoustic environment and results in a higher degree of immersion" (p. 93).

4.4.3 Full Ultra-Immersive Mask

Another interesting usage of technology can be found in the 'underwater-shooter' audio game Deep Sea (2010), where the player has to wear a modified gas mask in order to play the game. According to the developer's official website, the gas mask completely obscures the player's vision of the real world environment, while the virtual environment is presented aurally to the player through a set of headphones ("Deep Sea,"

[33] Examples of such head-tracking systems are the *Sony VPT* or the *Beyerdynamic Headzone* (Sony, n.d.; Beyerdynamic, n.d.).

n.d.). In addition, a microphone is implemented into the gas mask, which captures the real world breathing sounds of the player. These breathing sounds are then processed in real-time by the game system, and replayed to the player as if originating from the virtual environment. When processed by the game system, the breathing sounds are modified to sound as if the avatar is breathing through a diving tank under water—with a 'hissing' inhale sound, and a 'bubbling' exhale sound.[34] Furthermore, these underwater breathing sounds are intentionally exaggerated to mask the sound of enemies approaching (Renovitch, 2011). The player is consequently forced to breathe sparingly, in order to effectively locate where dangerous elements of the game are. The use of such an intimate factor as the player's breathing in the gameplay challenges, in combination with the near-claustrophobic feeling of having a gas mask strapped over the player's head (cloaking the sense of sight completely), makes the game Deep Sea possibly one of the most ultra-immersive audio game experiences created to date.

On an analytical level, the use of sound in Deep Sea (2010) becomes very interesting. The breathing sound is in the game perceived as emanating from the diegetic environment, since it represents the avatar's breathing process under water. As the game is highly ultra-immersive, however, the seemingly natural breathing sound of the avatar is not only produced by the virtual environment, but also directly generated as a reaction of the player's input.

Although the exact technological process behind the sound event is unknown to this study, there seems to be two possible processes that could have been adopted to generate this sound. The first would be that the sound of the player's breath is actually transduced as an audio signal, and processed in real-time by the game system to sound as underwater breathing. This would mean that the actual sound itself is originating from the real world environment, but at the moment of reproduction in the headphones it instead exists in the diegetic environment. The second possible process would be that the sound is pre-recorded by the designer, which upon control-related input from the player (a signal generated by the microphone in the mask) is being replayed by the game system as originating from the diegetic environment. Judging from a video of the game being played (Arnott, n.d.), it sounds like it could possibly be a combination of both of the processes; where the inhale sound is directly processed in real-time, and the exhale sound is a form of replayed, pre-recorded 'bubbly' sound. Regardless of

[34] For demonstration of Deep Sea being played, see http://vimeo.com/17299509 (Arnott, n.d).

production technique, the effect is that the sound event has an ultra-immersive quality, since it merges the player and the avatar into one unit. Consequently, this communicative relationship between the player and the virtual environment has clear transdiegetic properties, as the sound event not merely transgresses the diegetic border into the extradiegetic, but instead transgresses from the real world into the diegetic realm and vice versa.

On a side note, it could even be argued that transdiegetic communication is a vital factor in all ultra-immersive game experiences. In the light of the examples listed in this thesis (*Papa Sangre*, *BlindSide*, *Deep Sea* and *Zombies, Run!*), it seems that the core trait of any ultra-immersive digital game lies in its ability to effectively fuse the three game machinery entities together. Most importantly, it seems necessary that the Player and the Virtual Environment entities be merged into a single player–avatar unit, ineluctably involving the Game System both as an underlying technological link and as a control system with various degrees of mechanical intuitiveness to the player. All of the presented examples utilize transdiegetic communication between the parallel realities of the player–avatar unit, transferring information between the real world environment and the multiple diegetic levels of the virtual environment.

Returning to the analysis of the underwater-breathing sound, it has a direct usability function as it is caused reactively to a control-related input from the player. Furthermore, it also has an indirect proactive usability function, since the sound event masks with other sound events, making them difficult to hear if it is played too frequently. As a result of this, the player is urged to breathe sparingly in order to complete the gameplay challenges. The breathing sound also has both strong atmospheric as well as identifying functions, since it informs the player about the environment on both an informative and affective level. Additionally, as an orientation function, it distinguishes the event into two physically separate sources: (a) the mono representation of the inhale sound event, sounding as if originating from within the mask; and (b) the stereo representation of the exhale sound, sounding as if bubbles are emerging outside the diving mask on both sides of the avatar's head. Possessing properties of all of these functions, the sound event is used to both dynamise and optimise the gameplay experience.

In relation to Halliday's metafunctions, the sound is ideational by presenting information about the game world; it strengthens the interpersonal relationship between the player and avatar (virtual environment); and it is used as a textual form since it

provides a continuing, overarching texture to the situation by being a repetitive—yet rhythmically varied—sound event throughout the whole gameplay experience.

To conclude this chapter, it appears that through a creative use of technology, this one, seemingly simple sound event is in fact capable of simultaneously utilizing all of the functions and metafunctions presented in this thesis, as well as producing a sort of transdiegetic, looping feedback.

5 Summary: The Role of Sound in Audio Games

This final chapter of the study will take a look back on the main findings presented in the previous two chapters, and relate them together. In the introduction chapter, the word 'role' was defined according to Oxford Dictionaries as "the function assumed or part played by a person or thing in a particular situation" ("Role," n.d.). From this definition, it was concluded that a role is something one can *be* or *act* while performing a function within a context. Based on this, the thesis explored both the context as well as the functions of sound within audio games. This chapter will provide a more detailed definition of the two terms and present an analytical method that can be used to define an entity's role. In addition, the concept *Modes of Expression* will be introduced in this chapter, as it will be an important notion when defining the role of sound.

5.1 The Theatrical Interplay of the Game Machinery

The word 'role' itself is primarily derived from and related to a context of dramatic narrative enactments, such as theatrical dramas or, more recently, film productions. Although many other meanings of the word have been conceived over the centuries, and used in various social contexts, this thesis will here continue to draw parallels to the original narrative usage of the word.

Now, consider for a moment: if a digital game is metaphorically thought of as a sort of metatheatrical play, and sound as one of the actors participating in the narration—which types of characters could sound act as? In video games, sound would be but one of the actors, and the graphical representation would be another. They would both play very different roles, but through interaction on this hypothetical stage they work together to deliver the full narrative experience to the audience—the players. In this respect, audio games could be regarded as a monologue, where there is only one actor telling the whole narrative and describing the fictional world.

5.1.1 Modes of Expression

To think of digital games like a theatrical play would directly relate the integrated roles to the different modes of expression, which each has various traits and qualities. Any particular mode of expression has its own characterizing features, just as an actor has characterizing features that makes him or her suitable for certain roles.

The term ‘modes of expression’ is a more complex concept than merely a label of the direct channels to the senses, such as image or sound. In short, modes of expression relate to anything that expresses meaning of any kind. This could for example be colour, text, music, semantic use of voice or even non-semantic use of voice. Whenever combing many modes, into a so-called *multimodal expression*, the sum of them all might create entirely different meanings than when presented separately. Wingstedt (2008) emphasizes this by writing that “[e]ach mode individually bears meaning, but in the communicational process meaning typically emerges from the interweaving between and across modes within a multimodal system” (p. 28–29). Presenting happy music to a sad image, might for example implicate a sense of absurdity in the situation, which is not otherwise present in any of the two modes individually. In audio games, there are many variations of sound-related modes, coexisting and conjoined throughout the gameplay experience. These could for example be the different categories of the IEZA-model, each presenting various expressions of meaning to the player (e.g. music, ambience, effect sounds). In addition, the narrative story as a whole, or any semantic use of voice, could present other modes of expression within a game. Consequently, even without visual expression, audio games can still present a multimodal representation by combining different levels of aural meaning expressions.

Although sound can be regarded as an actor in a conceptual metatheatrical monologue, by utilizing the many modes of expression, sound may still present several ‘characters’ of meaning communicators. A real actor in a dramatic monologue may present different modes of expression simultaneously (e.g. gestures and speech), but would find it hard to act multiple roles at the same time and is often limited to taking turns between the characters. In digital games, however, numerous sound events and modes of aural expression are presented simultaneously. Depending on the level of perspective, this may call for a need to describe the role of each mode of expression individually—as opposed to talking about the role of the concept of sound itself. However, since individual modes can become a very detailed topic, such descriptions do not fit the limitations of this research.[35] Instead, this thesis will continue to think of sound as one actor, that although being present in multiple forms simultaneously in the same narrative, has its own unique set of characteristics and affordances.

[35] Wingstedt (2008) for example lists in detail six different narrative functions of only music, when played in a multimodal media production (p. 46).

5.2 The Characteristics and Affordances of Sound

Throughout this thesis, many characteristics and affordances of sound have been identified. The term *affordance* originated by Gibson's (1986/1979) study on visual perception.[36] In short, affordance relates to qualities and properties of an entity, and how these may offer certain action-related possibilities to another entity. In digital games, this could for example relate to a sound's affordance to provide a usability function, offering information that the player may use to make action related decisions in the game. This section will try to summarize the main findings of the study into an overview of the characteristics and affordances of sound in audio games.

5.2.1 The Affordances of the Game System in Audio Games

Although most hobbyist-developed audio games are still developed for computers, the mobile phone has been shown to be a popular platform for commercial audio games. Whereas video games become somewhat limited on mobile phones, because of their small screens and processing powers, audio games do not suffer the same extent of those restrictions. On the contrary, audio games on mobile phones can offer a more mobile game experience, and implement the real world environment in ways that are not possible for video games. In addition, modern mobile phones usually have some technological features that are hard to provide by computers. These features, such as the gyroscope sensor, can be implemented in the gameplay to provide ultra-immersive qualities to audio games.

As most audio games are intended to be played with headphones, the virtual environment effectively blocks out the external real world environment—assuming that the real world environment is not excessively noisy. In a video game, on the other hand, where the virtual environment is largely represented by graphics on a visual screen, the only equivalent to the immersive qualities of sensual representations would be to use a graphical display that covers the eyes (such as virtual reality glasses), as the virtual environment otherwise becomes dislodged from the player's immediate present. Audio games could therefore in general be regarded as having a higher potential of achieving a sense of immersion than normal video games. In combination with these immersive qualities when using headphones, it has also been shown that audio games can quite

[36] Gibson's original term has later been modified by others and used in various contexts, such as *Human–Computer Interactions* or *Social Semiotics.* The definition used in this thesis, however, is more general on purpose, and does not engage in discussions of perceived possibilities or other meaning differences.

easily gain ultra-immersive qualities by the implementation of some fairly simple narrative or technological features—qualities that are otherwise rather difficult to achieve in video games.

5.2.2 Sound as Audio and Sound as Sound

Within the context of a game system, sound may instead be referred to as *audio*. Although the terms 'sound' and 'audio' are similar in meaning and often confused, the definitions of the two terms are distinctive from each other, as audio usually denotes sound as an electrical process. Sound is what the listener would perceive, whereas audio is sound in its recorded or synthesised electrical or digital form. Albeit being a natural phenomenon, sound has the affordance of being transformed into audio, where it can be tamed and altered, and ultimately brought back into the aural world as sound, perceivable again by any who hears it. In fact, not only does sound afford to be turned into audio, sound *needs* to be captured, designed and reproduced in order to function as a coherent representation within a media. Moreover, since sound and audio can be designed, a listener can also evaluate a sound according to its aesthetical properties.

Multiple sounds have a tendency to mask each other, making them lesser than the total sum. This is not only an acoustical quality but also an electrical, digital and a psychoacoustical one. In acoustical, electrical and digital terms, multiple sound waves commonly interfere with each other when superimposed, causing both constructive and negative interference. In psychoacoustics, this phenomenon is even more complex, relating to both the binaural hearing process as well as the mental listening processing of sound. This has not explicitly been the focus of this study, as there already exists a rather comprehensive research field of acoustical sound propagation and psychoacoustical sound perception. Nevertheless, these are important traits of sound, and an important aspect to consider in any audio content media. In order to produce a satisfying sound environment, the sound designers need to create a balance of all presented sound events by mixing the audio according to many domains.

Sound also has the characteristic of being omnipresent. Sound is always an immaterial stream of events that cannot be paused, grasped and possessed. It is ever present in space, constantly flowing in time, and immerses any one who can hear it. Sound can often be the first thing that a person perceives before waking up and the last thing a person perceives before falling asleep. The hearing mechanism does not have a function to shut out sounds (like eyelids do for sight). However, when actively listening,

a person can psychoacoustically select which sounds to focus the attention to. Further on, a person's mind has the ability to create a mental imagery of an object or environment by hearing and listening to sounds.

5.2.3 Sound as Communication of Information and Meaning

The virtual environment of an audio game is primarily formed of sound, and in order to present a coherent message or meaning in relation to a listener, it needs to be shaped and designed intentionally by the game developers. In any digital game, the game designer will always be present as an ultimate meaning communicator, since all communication between the entities of the game machinery has been planned and designed beforehand by the developer. This relates to the affordance of sound as being a communicative vessel of information. Within audio games, sound is continuously used to provide the player with game-related information. By doing this, a sound event is also able to transcend its own border of diegesis, to indirectly affect events and actions on another.

Sound is always produced as an effect of its cause: an event of motion and life. And even though sound is only able to depict events, in doing so, sound may also indirectly provide information about objects or conditions involved in the event. Most environments filled with life and motion are also consequently filled with these causes of sound events. In an audio game, the designers have to reverse engineer this process, by filling the environment with sound in order to create a virtual world filled with life and objects. Since sound by affordance describes events of life and motion, sound greatly affects an environment's sense of presence and life in relation to a listener.

When a listener is placed within an environment, the physical interpersonal relationship between the listener and any sound event provides spatial information about the relative location of the event. Additionally, if used in certain ways, a sound event can also provide indirect information about objects and events not directly related to its own cause. In audio games, these affordances have generated a functional usage of sound as an orientational and navigational tool.

As two parts of a communicative act, the listener and the sound events work together in creating meaning about the environment. The listener draws from this meaning to produce a mental imagery of the information received. In addition to stimulating the listener's experience-based ability to identify and make sense of the surrounding

environment, sound can also be highly effective at suggesting and recalling emotions of a listener—giving it a great affectional narrative value.

Furthermore, apart from being inadvertently used by the listener to identify information on the cause of events, sound is also commonly structuralized to provide semantic information between communicators. Perhaps the most common of these semantic usages is through verbal language, but sound as a semantic vessel is also frequently utilized through other means; such as music, auditory signs or as a part of a multimodal meaning suggesting scenario.

5.3 Analysis of Sound as a Function and a Role

Throughout this research, extensive searches have been performed in order to find any type of definition of the word ‘role’ in relation to sound. However, not one proper definition on the ‘role of sound’ has yet been found—only loosely written usages of it. In this section, I will therefore try to further explain the definition of a role, and how it can be used in relation to sound. Further on, this section will also provide an analytical method for the definition and distinction of ‘role’ and ‘function’ of any entity in digital games.

5.3.1 A Problematic Matter of Definition

Although many people may have some intuitive understanding of the words ‘role’ and ‘function’, forming detailed definitions of the words and how they relate to each other has shown to be a tremendously complex task. What has become clear during this research is that the definitions and usages of the two words are very difficult to separate, even in other fields of research. Because of their similarity, they are constantly used interchangeably in various contexts, without any specification on which is what. This is, however, an understandable fact, since the two words share so many traits that, even when used incorrectly, the general meaning is often propagated anyway.

To begin with, one of the conflicting elements to this confusion may lie in the word class of the two words. When used in this manner, both of the words are nouns, even though ‘function’ commonly refers to an activity, whereas ‘role’ refers to an entity. There is nothing linguistically incorrect in describing activities by nouns—even though it may be more common to use verbs or adverbs for that purpose—but in this context it might make the distinction between the words ‘role’ and ‘function’ more difficult to discern. In addition, to define certain functions, it might sometimes be more reasonable

to use a noun instead of an adverb or adjective. For example, the phrase ‘sound has a *usability* function’ might provide a clearer meaning than to say ‘sound is functioning *useably*’ or ‘sound has a *useable* function’; especially since the term ‘usability’ is already a common term in the research field of human–computer interactions. Consequently, the usage of the word class noun (‘a function’), instead of the verbal form (‘to function’), is motivated and frequently used, but it highlights a part of the underlying problem of the repeated confusion of ‘function’ and ‘role’. When one says ‘to *have* a function’ and ‘to *act* a role’, the focus of ‘activity’ vs. ‘entity’ changes, as the ‘function’ in this case is perceived as a trait to possess and the ‘role’ is something one performs.

Another discordance between the words ‘function’ and ‘role’ may be found by looking at the origins of the word ‘function’. According to Oxford Dictionaries, ‘function’ is originally derived from the Latin words *fungi* or *functio*, meaning ‘to perform’ or ‘performance’ (“Function,” n.d.). Relating that original meaning of ‘function’ to the word ‘role’, the distinction between the two terms is again vague, since one might argue that the general function of a role is to perform something.

While trying to explain the definition of a role in the field of biomedical ontology, Arp & Smith (2008) make a fairly simple example by writing that a “heart has the function of pumping blood; but in certain circumstances that same heart can play the role of dinner for the lion” (p. 3). Further on, the authors conclude that “[r]oles are optional, and they often involve social ascription” (p. 3). To expand on this, we can also argue that the heart by itself has the function of pumping blood (or, more precisely, just *pumping*), but in the ‘social’ context of the body it has many different role relationships with the other entities of this context, such as the brain, veins, lungs, etc.

In order to properly transfer the analyzed entity from a condition where its functional affordances are clearly visible, into a situation where the entity instead possesses the properties of a ‘role’, we must firstly define more precisely what the entity of examination is, and secondly, define what the social context consists of and from the perspective of *whom* or *what* that entity becomes a role. Whenever the function of an entity relates to another entity, it can be claimed that a role relationship exists between the two. As exemplified above, the heart has the function of pumping blood, but it can also be said to have the function of pumping blood *to* the other entities of the body. Arguably, the definition of their respective roles in this context must therefore be found in what exists between the entities: in the interpersonal domain.

From this perspective, we can start to perceive two different analytical levels based on the definitions of the words 'function' and 'role', which will here be named as the *Functional* level and the *Interfunctional* level. The Functional level is the analytical level on which the individual entities, various modes of expression or any other asset included in a digital game or other social context can be analyzed according to their attributes and functional affordances. This is the level on which the details and conditions of both the context and the entity of examination are formed. The Interfunctional level,[37] on the other hand, is the analytical level on which the entity of examination can be analyzed according to the interpersonal attributes of its socially constructed roles in relation to the other entities of the specified context.

5.3.2 Establishing the Functional Level

In order to adopt the principles of the Functional and Interfunctional levels into this study, we are immediately faced with the difficulties of defining the entity of examination. A heart, as exemplified earlier, is a physical object, whereas sound is a much more intangible concept. Whenever analyzing such an abstract entity as sound, it quickly becomes very complex to discuss its 'functions' and 'roles', since the word 'sound' itself is not entirely easy to define. The basic definition is that sound is a phenomenon of nature, or as Oxford Dictionaries writes, "vibrations that travel through the air or another medium and can be heard when they reach a person's or animal's ear" ("Sound," n.d.). However, as has been described throughout this thesis, sound has many characteristics and properties other than just being wave propagations in a medium. And even though the entity of sound may have certain fairly evident functional affordances inherited by its nature (e.g. providing usability information to the player), its potential characteristics as a 'role' depends on how we choose to use the word 'sound'.

In this study, a wide-aiming perspective on the definition of sound within audio games has been adopted. Furthermore, this thesis has aimed to define the role of sound as a general concept. The entity of examination for the analysis in this research is therefore 'sound as a general concept', possessing all of the previously stated characteristics and affordances. In audio games, this would relate to the whole stream

[37] The analytical perspective of this level is focussed on the relationships among and between the entities of the social context, hence, the use of the prefix *Inter-*. Oxford Dictionaries: "Inter-: *prefix* 1, between; among. *prefix 2*, mutually*;* reciprocally" ("Inter-," n.d).

of designed sound events, originating from within the virtual environment, passing through the game system, and ultimately perceived and mentally processed by the player.

After the entity of examination has been defined, the context of the Functional level needs to be defined as well. Throughout the study, this context has been the notion of the *Audio Game Machinery*, comprised of the three entities the *Game System*, the *Player* and the *Virtual Environment*. This context has been thoroughly explored and the functional purposes between the entity of examination (sound) and the other entities of the functional level have been established earlier in the thesis. In establishing the functional level, however, we see that many of the included entities share a relationship together *through* the various functions (e.g. a heart has the function of pumping blood to the other organs). Is it therefore possible to analyze the entities and their functions individually without also closely examining their relationships on the same analytical level? Reflecting on this for a moment brings the thoughts to the old philosophical question: if a tree falls in a forest and no one is around to hear it, does it still make a sound? The term functional *affordance* again becomes relevant here, as it focuses on *possible* functional offerings of the entity, without necessarily specifying the social context. For example, a hammer offers the affordance of *hammering*. Information on whoever decides to use it, or for what, has lesser relevance in this perspective. On the functional level, it is therefore possible to individually analyze the functional affordances of an entity, even if the context decrees that other entities are indirectly involved in the function—as their relationships are not the main focus of the functional level's perspective. Parallels can be drawn to Halliday's (1978) previously explored ideational and interpersonal metafunctions: on an ideational level, the communicator's potential in making meaning is individually analyzed according to his or hers subjective perception of the world, whereas on an interpersonal level it is the *relationship* between the communicator and the world that is being analyzed.[38]

A simple example of an analytic statement on the functional level could be: sound in an audio game has the functional affordance of providing information about the virtual environment to the player via the game system. In this example, the context is set (an audio game); the entity of examination is defined (sound); the entity's functional

[38] For further descriptions of Halliday's metafunctions, see section *4.1 – Metafunctions of Communication*.

affordance is specified (providing information); other entities of the context are present (the player, the virtual environment and the game system); and also, relationships between the entity of examination and the other entities can be established (sound provides information *about* the virtual environment *to* the player *via* the game system). This is however only one single statement—one small part of this specific context—and the more detailed the Functional level as a whole is analyzed, the more prepared a study will be to make an analysis on the Interfunctional level.

As a social situation, involving relationships among entities and with information being transferred between the participators of the context, the context of the audio game machinery fulfills the prerequisites of being analyzed on an Interfunctional level. By taking all of the known information on the Functional level, and shifting the perspective to focus on examining the characteristics, traits and affordances of the *relationships* between the entities of the context, the Interfunctional level aims to accurately describe the social properties of the entity of examination—which will define its roles.

5.4 The Roles of Sound in Audio Games

The general conclusion of this research is that an entity can only have one role relationship towards another entity within a specific social context. An entity should therefore not be labelled with multiple roles just to describe various features of the entity. The entity of examination may, however, have one role relationship with one entity present in the social context, and another role relationship with another entity of that same (or other) social context. It is therefore possible for the entity of examination to act multiple roles at the same time, but only if these roles are shown in relation to other entities of the social context individually. Any role can then be described to have certain features, which will define and characterize the role from the perspective of its role related other.

On the functional level, the examined entity in this study is the concept of sound, as is present with all of its characteristics, affordances and functions in its natural condition within the social context of the audio game machinery. In this context, there are three other main entities present—the game system, the player and the virtual environment. According to this perspective, this means that sound in audio games acts three main roles; one role in relation to the game system; one role in relation to the player; and one role in relation to the virtual environment.

5.4.1 The Role of Sound in Relation to the Game System

Sound is the communicative tool of the game system in audio games. With sound, the game system is able to represent the virtual world and the gameplay challenges. By examining the interfunctional role relationship between sound and the game system, it can be discerned that although sound is the medium used for the communicative act by the game system, the game system is the medium used by sound to propagate to a listener. It could be claimed that the game system has a dominating status over the sound, since the game system orders and controls which sound should be communicated, and subsequently carries the sound through its system chain until it reaches the player. However, as was concluded earlier, the game designer can be regarded as the ultimate communicator in a game. This implies that the designer uses sound to represent the virtual environment, and subsequently orders and controls the game system to make sure that sound is communicated properly. The two entities, sound and game system, could therefore be regarded to share a synergistic relationship, where anomalies in one imposes distortion of meaning in the other.

5.4.2 The Role of Sound in Relation to the Player

In audio games, sound acts a critical role in the gameplay experience of the player, since sound provides the only representation of the virtual environment. Their interfunctional relationship therefore indicates that sound has a superior status over the player. Sound is used to present the virtual environment, and guide the player within it. As a tool of narration, sound is used to describe the game plot to the player, and inform the player of the events that occur within the story. The player is also dependent on sound to receive feedback of its own interactions within the game. Although these interactions are necessary for the game to progress, the player does not communicate to the sound events directly, but indirectly via the game system and the virtual environment. The relationship between sound and player is therefore not as synergistic as the relationship between sound and game system, even though the player has a vital part of the game machinery as a whole, and therefore also indirectly to the sounds themselves.

5.4.3 The Role of Sound in Relation to the Virtual Environment

In a virtual world, the simulated and pre-designed reality *is* the reality of that universe, and as with any universe, it always comes with a set of physical laws—even if they are only fictional simulations of our own. This means that, on a diegetic level in three-dimensional audio games, sound will be a simulated reaction to its cause. Sound can

therefore be regarded as a natural phenomenon, bound to the rules of the virtual environment. Even on the extradiegetic level, and with synthesized sounds without any logical causal reference in the game universe, sound still originates from within the virtual environment and is thus subordinate to it. Only by including the perspective of the designer will sound in audio games be dominant to the virtual environment, as the designer needs to reverse the process of cause and effect in order to create a fictional physical world through sonic means alone.

When a game is being played, the virtual environment is primary. It is a multileveled universe created by the designer; realised through the game system; inhabited by the avatar along with other entities of the narration; and ultimately experienced and lived by the player. Without the virtual environment, sound has no host universe to be mediated within. However, without sound, the virtual environment in audio games will be nothing but the blueprint to a design; a void without time or space—a condition suspended *in vacuo* before the big bang. Consequently, the interfunctional relationship between the two holds the virtual existence together like an adhesive bond. Sound and the virtual environment both create each other and are being mutually created by the other at the same time. The role of sound in relation to an aural virtual environment is therefore an ambiguous one, acting simultaneously as both predominant and inferior: forever ruled by the environment, yet always essential to its existence.

5.5 A Final Note

Conclusively, all of the functions, purposes, affordances and social attributes of sound make it a powerful entity, capable of independently providing a full gameplay experience within a digital game machinery. Naturally, sound has an extremely important part to play in the game form of audio games. All of its role relationships point to either an equal or a dominant status within the audio game machinery. In contrast, this superior status of sound is not commonly present in video games, where sound shares the role of being the communicative representation of the narrative and virtual environment with the graphical entity. Although the many functionalities of sound may be similar in audio games and video games, the roles of sound are very different between the two game forms. The definition of the roles of sound in video games, however, will be left for another study, as they have not been the focus of this research.

6 Conclusions

When I started preparing for this research I saw the need to bring forth certain aspects of sound functionalities within audio games. I did not then know, however, that the project would grow to this size. Continuously throughout the research, there has been an accumulating effect regarding the need to read. It would seem, that the more I read, the more I realised had to be read, as one concept or phenomena always needed to be explained by another. This process has still not stopped or slowed down, rather the opposite—which may also be befitting to the ever-progressing nature of digital games.

6.1 Critical Reflections on the Methodology

Initially, one of the proposed methods of this thesis was to recreate a modified version of Jørgensen's methodological approach in her study of game sound functionality in video games, but applied instead to audio games to investigate how the results might differ. This was supposed to have been carried out by making an in-depth analysis of the audio game Papa Sangre (2011), as well as arranging a series of listening-tests followed by interviews. Unfortunately, however, Papa Sangre was removed from the App Store due to maintenance improvements by the developers, and stayed like that for the whole duration of this study. In the end, I believe that this turned out to be for the best, since it made me re-think the direction of the thesis.

Although many other methods and approaches were also considered, I believe that the one that was finally used proved to be efficient for the purpose of the study. The major limitation of this research is that it was set to investigate the field from a very wide perspective. Although I believe that this was necessary to achieve the results that were presented in this essay, further research with a more focussed perspective on all of the identified issues will surely offer a better understanding of the research field.

Any research based on the contemporary content of a field will be limited to what is present at the time. Since the field of audio games is still in an early experimental paradigm phase, it therefore makes it difficult to examine *possible* functions of sound within audio games if they have not yet been realised. It is my belief that as the art world of audio games matures, we will see many new and interesting ways of using sounds within the games. Many artists use limitations to enhance creativity, and the limitation of using only sounds will most likely force this creativity among audio game developers to heights that are still unknown to us.

6.2 Critical Reflection on the Results

The purpose of this study was to help establish a theoretical ground and stimulate further research in the field of audio games. Throughout the thesis, many of the issues relating to sound in audio games have been presented and addressed. In doing this, many previously unaddressed issues have also been identified and brought up, which if researched further would present even more knowledge to the field. I therefore believe that the purpose of the thesis has been achieved.

As far as possible in a study like this, I consider that the research questions have been answered fairly extensively. However, being a very abstract object of research, there can be many ways of concretizing an answer. Further research based on other theoretical perspectives is therefore likely to find alternative answers to these addressed issues.

Section *3.1 – Conditions of the Game System & the Development Scene*, was one of the more difficult to form, since I found little or no formal research done previously on the issues addressed there. Although I tried to be as thorough as possible in the data gathering for that section, it was difficult to implement the findings in a formal text. In addition, it was difficult to keep track of what information was learned where, and what needed to be referenced or what could pass as general information or assumptions. In the end, although the section provides an important backdrop to the following chapters, that section was the one that had the least connection to the research questions. I therefore believe that the potential validity value of that section does not negatively affect the results of this research.

6.3 Further research

Originally, this research project aimed to create a graphical framework of how the currently established theories of game sound functionalities relate to both each other and to the games themselves. This idea, however, was quickly abandoned when the magnitude of theoretical exploration that such a research would have to encompass became obvious. Since there are many different theories on game sound functionality, it would be very difficult to summarize them as they all focus on these issues from different perspectives and use slightly different terminologies. Still, I do hope that future studies will explore this notion further, to form a more comprehensive overview of the different theories related to game audio.

As I wrote in segment *3.3.2 – Balancing the Sounds*, I warmly welcome more research in the area of soundscape fidelity and listener attentiveness, especially to conceptualize listening modes adapted to the playing of digital games. Further research of the theories of listening presented by Chion, Smalley, Schaeffer and other authors of that field, in combination with the notion of soundscape fidelity and an interactive audio environment, will surely bring more extensive knowledge about how players listen while playing a digital game and how this can be utilized by game developers.

Furthermore, in segment *4.3.2 – Direct and Indirect Functional Usages* this study identified two new factors to consider when analyzing game audio functionality. The theory presented there was however only briefly explored and I therefore welcome further research on that topic in order to reject, confirm or expand the theory. I also hope that future studies will test the theory of the Functional and Interfunctional analytical levels, as identified in section *5.3 – Analysis of Sound as a Function and a Role*.

On the linguistic side, I believe that the terms 'function' and 'role' should be researched further in order to be more defined and distinguished from each other. This thesis has showed that the distinction between the two words is important, as it forces the researcher to define specifically what the object of examination really is. I therefore suggest that further studies should reflect upon the definitions of the two words, and the potential meaning differences they imply. In order to create a more comprehensive theoretical field of sound and audio in digital games, I believe that clarity and precision in defining the involved entities and their various relationships are essential. For example, if one wishes to study the 'role of sound in relation *to* audio games', audio games is no longer the context, but an entity in a larger context (e.g. digital games). It might also be needed to expand on the notion of the digital game machinery, to potentially identify and include sub-entities, such as narrative or the designer. Furthermore, I would also recommend future studies to create graphical models (such as the IEZA-model) that overviews all of the entites and sub-entities of the digital game machinery, with all of their functional affordances and interfunctional relationships mapped out.

Lastly, the field of game audio—especially in audio games—is still a highly underexplored field of research. Further research on all of the issues relating to sound in digital games is therefore vital, in order for us to better understand the creatively challenging, functionally powerful and emotionally effective art of designing sound.

7 Bibliography

Literature and Articles

Arp, Robert & Smith, Barry (2008). *Function, Role, and Disposition in Basic Formal Ontology*. Retrieved January 22, 2013 from http://nemo.nic.uoregon.edu/wiki/images/7/7a/Arp_et_al-NPre2008_Function-Role-Disposition_in_BFO.pdf

Barbosa, S. D. J., de Souza, C. S. & de Paula, M. G. (2003). The Semiotic Engineering Use of Models for Supporting Reflection-in-Action. In: J. A. Jacko & C. Stephanidis (Eds.): *Human–Computer Interaction: Theory and Practise (Part 1)*. (pp. 18–21). New Jersey: Lawrence Erlbaum Associates, Inc. Retrieved from http://books.google.se

Beck, Alan (1998). *Point-of-listening in Radio Plays.* Retrieved December 31, 2012 from http://www.kent.ac.uk/arts/sound-journal/beck981.html

Björk, S. & Holopainen, J. (2004). *Patterns in Game Design* [Adobe Digital Editions]. Retrieved from http://www.ebrary.com

Chion, Michel (1994). *Audio-Vision: Sound on Screen.* (C. Gorbman, Trans.) New York: Colombia University Press

Ferrington, Gary (1993). Audio Design: Creating Multi-Sensory Images for the Mind. *Journal of Visual Literacy, 14*(1), pp. 61–67. Retrieved from http://www.ohio.edu/visualliteracy/v14_1.html

Forrester, Michael A. (2002). *Psychology of the Image* [Adobe Digital Editions]. Retrieved from http://www.ebrary.com

Friberg, J. & Gärdenfors, D. (2004). Audio Games: New Perspectives on Game Audio. *Proceedings of the 2004 ACM SIGCHI Conference on Advances in Computer Entertainment Technology*, pp. 148‑154. Singapore: ACM Press.

Gibson, James J. (1986). *The Ecological Approach to Visual Perception.* New Jersey: Lawrence Erlbaum Associates, Inc. (Original work published 1979) Retrieved from http://books.google.se

Grimshaw, Mark (2007). *Sound and Immersion in the First-Person Shooter*. Retrieved from http://hdl.handle.net/2436/35995

Halliday, Michael A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold Ltd.

Huiberts, Sander (2010). *Captivating Sound: The Role of Audio for Immersion in Computer Games* (Doctoral thesis). Available from British Library Electronic Theses Online System (uk.bl.ethos.529027)[Low-quality Scanned Version], though retrieved October 30, 2012 from http://captivatingsound.com/phd-thesis-captivating-sound-the-role-of-audio-for-immersion-in-games/ [High-quality Digital Version]

Izhaki, Roey (2010). *Mixing Audio – Concepts, Practises and Tools*. Oxford, UK: Elsevier Ltd.

Jørgensen, Kristine (2009). *A Comprehensive Study of Sound in Computer Games: How Audio Affects Player Action*. Wales: The Edwin Mellen Press, Ltd.

Malina, Debra (2002). *Breaking the Frame: Metalepsis and the Construction of the Subject*. Columbus: The Ohio State University. Retrieved from http://books.google.se

Møller, H., Sørensen, M. F., Hammershøi, D., & Jensen, C. B. (1995). Head-Related Transfer Functions of Human Subjects. *Journal of the Audio Engineering Society, 43*(5), pp. 300–321. Retrieved from http://www.aes.org

Rodaway, Paul (1994). *Sensuous Geographies: Body, Sense and Place* [Google Books version]. Retrieved from http://books.google.se

Röber, N. & Masuch, M. (2005). Leaving the Screen: New Perspectives in Audio-Only Gaming. *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*, pp. 92–98. Retrieved January 24, 2013 from http://citeseerx.ist.psu.edu

Schafer, Murray R. (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books (Original work published 1977)

Smalley, Denis (1996). The Listening Imagination: Listening in the Electroacoustic Era. *Contemporary Music Review*, *13*(2), pp. 77–107. doi:10.1080/07494469600640071

Sikorski, Marcin (2003). Measuring the Immeasurable: System Usability, User Satisfaction and Quality Management. In: J.A. Jacko & C. Stephanidis (Eds.): *Human–Computer Interaction: Theory and Practise (Part 1)*. (pp. 411–415). New Jersey: Lawrence Erlbaum Associates, Inc. Retrieved from http://books.google.se

Sonnenschein, David (2001). *Sound Design: The Expressive Power of Music, Voice and Sound Effects in Cinema.* Los Angeles: Michael Wiese Production. Retrieved from http://books.google.se

Van Leeuwen, Theo (1999). *Speech, Music, Sound*. London: Macmillan Press Ltd

Wingstedt, Johnny (2008). *Making Music Mean: On Functions of, and Knowledge about, Narrative Music in Multimedia* (Doctoral dissertation). Sweden: Luleå University of Technology, Department of Music and Media. Retrieved January 20, 2013 from http://epubl.ltu.se/1402-1544/2008/43/LTU-DT-0843-SE.pdf

Games

BlindSide (1.02) [Digital game for iPhone]. (2012). USA: Epicycle llc. Official website: http://www.blindsidegame.com/

Deep Sea (2.0) [Computer software]. (2010) USA: Wraughk. Official website: http://wraughk.com/deepsea/

Papa Sangre (1.1.3) [Digital game for iPhone]. (2011). London: Somethin' Else. Official website: http://www.papasangre.com/

Swamp (2.8) [Computer software]. (2012). USA: Aprone. Retrieved from http://www.kaldobsky.com/audiogames/

Zombies, Run! (1.5.1) [Digital game for iPhone/Android /Windows phone]. (2012). London: Six To One. Official website: https://www.zombiesrungame.com/

Internet Sources

About (n.d.). In About | Papa Sangre. Retrieved December 17, 2012 from http://www.papasangre.com/

Alexander, Jem (2012, August 28). *Version 1.4 for iOS coming soon.* Retrieved January 19, 2013 from http://blog.zombiesrungame.com/post/30384677108/version-1-4-for-ios-coming-soon

Arnott, Robin (n.d.). *Deep Sea (v 2.0. Audio)* [Video file]. Retrieved January 24, 2013 from http://vimeo.com/17299509

BAFTA Guru (2012, September 03). *Nick Ryan (The Night Jar, Papa Sangre) : Sound Design Mini Masterclass.* Retrieved January 17, 2013 from http://www.youtube.com/watch?v=Roa1wI3MS5M

Beyerdynamic (n.d.). *Beyerdynamic Headzone Base Station H 1*. Retrieved January 24, 2013 from http://europe.beyerdynamic.com/shop/headzone-headphone-surround-system.html

First Post! (2010, March 04). In Blog | Papa Sangre. Retrieved December 17, 2012 from http://www.papasangre.com/blog/#post-25

Fiction (n.d.). In Oxford Dictionaries. Retrieved January 20, 2013 from http://oxforddictionaries.com/definition/english/fiction

Function (n.d.). In Oxford Dictionaries. Retrieved January 11, 2013 from http://oxforddictionaries.com/definition/english/function

Inaudible (n.d.). In Collins English Dictionary. Retrieved January 19, 2013 from http://www.collinsdictionary.com/dictionary/english/inaudible

Inter- (n.d.). In Oxford Dictionaries. Retrieved January 23, 2013 from http://oxforddictionaries.com/definition/english/inter--

Invisible (n.d.). In Collins English Dictionary. Retrieved January 19, 2013 from http://www.collinsdictionary.com/dictionary/english/invisible

Role (n.d.). In Oxford Dictionaries. Retrieved January 11, 2013 from http://oxforddictionaries.com/definition/english/role

Renovitch, James (2011, April 15). *Sinking Deeper.* Retrieved January 24, 2013 from http://www.austinchronicle.com/screens/2011-04-15/sinking-deeper/

Sony (n.d.). *Sony Global – VPT(Virtualphones Technology).* Retrieved January 24, 2013 from http://www.sony.net/Products/vpt/index.html

Sound (n.d.). In Oxford Dictionaries. Retrieved January 22, 2013 from http://oxforddictionaries.com/definition/english/sound

TheUnixCube (2010, December 28). *Papa Sangre Review for iPod Touch 4G + iPhone 4 App Game Test 1st LvL, Music & Sound - Dec 28, 2010* [Video file]. Retrieved January 15, 2013 from http://www.youtube.com/watch?v=yObFPLdDToo

Top 200 (Paid) in Healthcare & Fitness for iPhone (2012, December 15). In AppShopper. [Top list] Retrieved December 15, 2012 from http://appshopper.com/bestsellers/healthcare-fitness/paid/?device=iphone

van Tol, R. & Huiberts, S. (2008). *IEZA: A Framework For Game Audio.* Retrieved January 21, 2013 from http://www.gamasutra.com/view/feature/3509/ieza_a_framework_for_game_audio.php

Watson, Krystal (2012, March 07). *Zombies, Run!* Retrieved January 19, 2013 from http://www.applevis.com/apps/ios/health-and-fitness/zombies-run

Appendix A - Transcript of ‘A Fluttery Watery Thing’

Gameplay video of the first level of the game Papa Sangre

TheUnixCube. (2010, December 28). *Papa Sangre Review for iPod Touch 4G + iPhone 4 App Game Test 1st LvL, Music & Sound - Dec 28, 2010* [Video file]. Retrieved January 15, 2013 from http://www.youtube.com/watch?v=yObFPLdDToo

Time cue: 2:48

> Here you are in the land of the dead—the realm ruled by Papa Sangre. You are a brave soul, here to rescue a soul in peril, no? In this underworld it is pitch-dark. You cannot see a thing. You cannot even see me, a fluttery watery thing—here to help you. But you can listen and move. I will tell you how. First you must learn how to walk in this body. Next you must learn how to see with your ears. You will need these powers to save the soul in peril and make your way to the light. First you must learn how to walk in this body. Walk with your thumbs and listen to your footsteps. Left. Right. Left… [Player walking] Yes, you are walking… [Player walking] A little more…

Appendix B - Summary of Forum Discussions

B.1 – CAE_Jones (2012-12-11 03:30:22)

User name: CAE_Jones

Forum thread name: *Research on the Role of Sound in Audio Games*

"Re: 1. Most audio game developers are individuals with limited funds and other jobs that force game development to take hobby status. So yes, the Papa Sangre team is a huge exception to the norm, but I also think they came from a purely mainstream perspective and weren't familiar with the existing audio games community, which has its advantages and disadvantages (I'm not sure; trying to remember from the thekillscreen article). Most of the games I have in the works would be hard to pull off on my own, and my experience with trying to recruit people and manage a multi-person project tell me that I am a terrible manager of volunteers, and would do better if I could afford to hire a team. (One or two assistant programmers, a manager who can handle people better than me, a dedicated sound person and lots of voiceactors would help tons.). Entombed sold 500 copies, if I remember correctly, which is taken as the record for audio games intended for the audio games market. I don't know how well Papa Sangre and the Nightjar have done, but I suspect that they've probably sold better than Entombed. There seems to be a hint that mainstream developers are getting more experimental with interface and story/character immersion, so I wouldn't be surprised to see more games like Papa Sangre in the future."

Retrieved January 20, from:
http://forum.audiogames.net/viewtopic.php?id=9410

B.2 – SLJ (2012-02-28 10:05:07)

User name: SLJ

Forum thread name: *new iOS release, zombies, run!*

"Hi.

Yeah this game sounds totally amazing. Well I would like to run, but running as a blind guy in a city full of traphic while wearing headphones, no way!  I hope you can cheat a bit by running in your own house or just step like you run. I've really looked forward to this game and I can't wait to check it out."

Retrieved January 19, 2013 from:

http://forum.audiogames.net/viewtopic.php?id=7450

"ALL RESEARCH INTO SOUND MUST CONCLUDE WITH SILENCE—NOT THE SILENCE OF NEGATIVE VACUUM, BUT THE POSITIVE SILENCE OF PERFECTION AND FULFILMENT. THUS, JUST AS MAN STRIVES FOR PERFECTION, ALL SOUND ASPIRES TO THE CONDITION OF SILENCE."

– MURRAY R. SCHAFER

DALARNA UNIVERSITY COLLEGE
RICKARD ÅSÉN