APPROACHES TO AUDITORY INTERFACE DESIGN – LESSONS FROM COMPUTER GAMES

Graeme W. Coleman\textsuperscript{1}, Chris Hand\textsuperscript{2}, Catriona Macaulay\textsuperscript{2}, Alan F. Newell\textsuperscript{1}

\textsuperscript{1}Division of Applied Computing, 
\textsuperscript{2}Interaction Design Lab, 
University of Dundee, 
Dundee, 
Scotland, DD1 4HN. 
{gcoleman, chand, catriona, afn}@computing.dundee.ac.uk

ABSTRACT

The computer game has begun to establish itself within the wider entertainment industry, and has thus attracted considerable interest from more general interaction designers. However, while computer game audio has become increasingly sophisticated, it remains a discipline largely overlooked by the research community. We begin by outlining similarities between each discipline, highlighting those which we believe provide interesting opportunities for designers of auditory interfaces. We also suggest that, through an understanding of the everyday practices of computer game sound designers and their colleagues within the industry, the process of sound design for alternative forms of interfaces can be considerably informed. To discover and understand some of these practices, we present our experiences conducting a field study using ethnographic methods with a major UK-based computer game developer. We highlight discoveries which we believe are pertinent for the design of auditory interfaces and thus merit further research.

Our study forms part of our wider research to develop a grounded theory (i.e. a theory conceived via the data collected during the field study) to understand the reality of sound design within the computer games industry, relationships to the design of more general interfaces and thus how we approach the design of contemporary auditory interfaces.

1. INTRODUCTION

At ICAD '04, Leplâtre and McGregor acknowledged that, despite twenty years of auditory display research, computers have largely remained silent with the exception of games and multimedia design [1]. Similarly, in 1997, Gaver suggested: “In general, research on auditory interfaces could learn a great deal about effective sound design from multimedia and games work” [2]. Indeed within a similar timeframe, the computer game soundtrack has evolved dramatically from the crude sounds of Pong and Defender. Poole [3] suggests that many hours are devoted “to make the whole audiovisual experience as immersive and (deceptively) ‘authentic’ as possible”. However, while more general studies have been carried out to examine the specific features of computer games which could help to improve the usability of applications in general (such as user performance and satisfaction) [4, 5], the specific characteristics of computer game audio have largely been overlooked.

Furthermore, in addition to understanding the use of sound in computer games, a further concern is to understand approaches to sound design. To facilitate the immersive and authentic nature of a computer game, sound design must form a significant part of the development lifecycle. Indeed, as we shall see, many major developers operate full time audio departments who are involved during the entire game development process. This poses a number of interesting questions: for example, between the initial concept and the final production of a computer game, at which point does the design of sound become an issue? Why are specific choices for sound made? Consequently, where does the sound designer “fit” into the overall design process? What influences do they have? An understanding of such issues, which can be achieved via the studying of work practices within the industry, has promising consequences for auditory interface research. We propose that the result is a situation in which sound design forms a greater role in the overall design process, resulting in a more appropriate interface. In the next few sections, we open this discussion by providing further suggestions as to why computer game audio is a useful discipline to investigate. We further outline the motivation behind our field study and discuss the methods used in the work presented, before reporting our experiences with a major UK computer games developer. Finally, based on our experiences, we consider some potential areas for future research.

2. THE CASE FOR COMPUTER GAMES

There is no doubt that the computer games industry is growing in stature, dispelling the popular myth that games are a product of the lone “bedroom programmer” (as was the case for many successful games sold during the early 1980s). In financial terms, a recent worldwide forecast suggests worldwide game industry and software revenue shall continue to grow from $23.2 billion in 2003 to $33.4 billion in 2008, excluding the large market for rentals, used games, accessories, books, magazines and many online games [6]. Of course, while these figures highlight the almost ubiquitous nature of the computer game, they do not identify similarities between the design of games and more general interfaces. Indeed, concentrating on computer game audio, we propose similarities between computer game sound design and the design of auditory displays.

Rollings and Adams raise some important issues: “Even though sound is often in third place after the visual and interactive elements, the fact that many games are unplayable without it clearly indicates the importance of sound” [7]. Firstly, while it is arguable that this correlates with a more general human attitude (at least in the West) to our auditory versus visual senses, the issue of sound residing in “third
places” resonates with research into more general user interfaces, where interaction has predominately been (and still is) determined by visual cues. It is interesting that in neither a computer game nor more general human-computer interaction does audio play the most prominent role. Yet (unlike more general applications) the idea of a contemporary commercially available silent computer game is almost unheard of.

Furthermore, with regards to game audio, Rollings and Adams concede that “games are unplayable without it” [7]. Although disputing whether the use of audio improves game playability falls beyond the scope of this paper, it is conceivable that the increasingly advanced audio capabilities of games consoles, and the vast majority of games which take advantage of these technologies, is an attempt to validate this statement. Similarly, whilst it is not proposed that user interfaces are completely unusable without sound in the same way that games are unplayable without sound, it is arguable in a broader sense that a suitable auditory interface is one in which navigation is vastly reduced in the absence of audio.

So what can computer games offer? Poole [3] suggests that sound is used within games for both functional purposes (“any sound can become a clue, a spur to action”) and to enhance the player’s involvement. Describing “Silent Hill”, a horror game developed by Konami for the Playstation, Poole describes how the player’s character is provided with a radio which initially appears to be broken, but emits a nerve-fraying noise whenever an enemy approaches — “The evocation of fear is deliciously heightened by this aural sign, as you run around panicking when the alarm goes off, not knowing from which direction the beast is going to approach through the omnipresent fog” [3]. The use of audio is more than functional here; the player is immersed in the game environment. As Poole suggests: “The reason sound design is important in videogames is quite simple: if a laser makes a pleasing, fizzy hum, and if an exploding enemy makes a particularly satisfying boom, then the game is just more fun to play” [3]. Hence, designers of auditory interfaces could gain knowledge from the way games are “more playable” through the use of sound.

Naturally, there are significant differences between computer games and more general applications to consider. For example, computer games, by their very nature, are “playable” through the use of sound. If a laser makes a pleasing, fizzy hum, and if an exploding enemy makes a particularly satisfying boom, then the game is just more fun to play” [3]. Hence, designers of auditory interfaces could gain knowledge from the way games are “more playable” through the use of sound.

is, we concentrate on the everyday social, lived, reality of sound design within a games development company, and the issues the sound designer and their colleagues must face during this process. It is hoped that an understanding of these issues shall inform designers as to how to approach the design of auditory interfaces, and researchers as to consider possible research paths.

3. METHODS

3.1. Motivation for Ethnographic Research

Strauss and Corbin state of qualitative researchers: “[They] reject statistical and other quantitative methods as yielding shallow or completely misleading information. They believe that to understand cultural values and social behavior requires interviewing or intensive field observation, with these being the only methods of data collection sensitive enough to capture the nuances of human living” [10]. Whilst it is beyond the scope of this paper to provide a comprehensive review of qualitative research methods, and their comparisons to quantitative methods, the rewards for using such methods can be extremely valuable.

Contemporary ethnographic studies attempt to understand general human behavior within particular contexts, to understand why participants act in a certain manner under which circumstances [11], and to understand how cultural meanings can be derived by the way participants “collide and mix” in changing situations in order to find solutions to problems [12].

The basis of ethnographic research involves fieldwork, i.e. becoming intimately familiar with the participants and their social activities for research purposes. Normally, ethnographers will carry out their research within the natural setting of the participants (the “field”) using principle methods such as observation, interviews, “desk research” (research carried out using documentation and records kept by participants) and surveys. Participant observation, whereby researchers directly take part in the culture and lives of the participants while keeping a professional distance to allow adequate observation and data collection [11], is a useful method to adopt. The boundaries between “participating” and “observing” are somewhat narrow, although Wolcott suggests that the key to successful participant observation is “to participate more and to play the role of the aloof observer less” [13].

Researchers within human-computer interaction have suggested that traditional experimental psychology cannot provide a broad enough conceptual frame to understand human activities nor a set of useful research methods for the design of computer-based information systems [14]. Therefore, ethnographic methods have been suggested to provide an adequate understanding of the nature of work underpinning the creation of interactive systems [15]. Discussing qualitative research in general, Miles and Huberman suggest that unexpected discoveries can be found, thus helping researchers to widen their analysis beyond initial preconceptions and frameworks, further suggesting: “Words, especially when they are organized into incidents or stories, have a concrete, vivid, meaningful flavor that often proves far more convincing to a reader – another researcher, a policymaker, a practitioner – than pages of numbers” [16]. Within human-computer interaction, Burke and Kirk suggest that valuable data can be discovered through ethnography which may never have been discovered.

---

1 There are exceptions to this rule, of course, e.g. “Pop Idol” by Codemasters, in which the player uses the game controller or “dance mat” to ensure the game character sings in tune and thus progresses through each level.
through the use of walkthrough tests by the interface designers [17]. Thus, drawing on the experiences of those working with audio in other domains, it is arguable that some of the “incidents or stories” which arise can help to inform designers of auditory interfaces and to offer several possible avenues for future research.

3.2. Grounded Theory

This paper represents part of a wider grounded theory-oriented field study of sound design in the industry. A complete theory is not developed here; however, we shall attempt to summarize the key points as a basis for the rest of this paper.

Briefly, grounded theory was first described by Glaser and Strauss in 1967 as “the discovery of theory from data systematically obtained from social research” [18]. Strauss and Corbin state that “theory denotes a set of well developed categories (e.g. themes, concepts) that are systematically interrelated through statements of relationship to form a theoretical framework that explains some relevant social, psychological, educational, nursing, or other phenomenon” [19]. This theory is generated from research “grounded” in data collected qualitatively rather than through quantitative forms of analysis.

Since the majority of data collected remains unstructured, Strauss and Corbin illustrate a process of analysis known as “coding” in which data are simultaneously fractured, conceptualized and integrated to form a theory [19]. Three different types of coding are suggested: open coding, axial coding and selective coding. Briefly, open coding is a process in which abstract concepts (the “building blocks” of theory) are identified and categorized to indicate “what is going on” within the phenomena being studied. Axial coding is a process in which similar categories are developed and linked, while selective coding is the process of integrating and refining a theory.

Sarker et al [19] suggest that grounded theory is valuable for three reasons. Firstly, developing a theory prevents researchers imposing theories from a related substantive area which does not actually match patterns in the data. Secondly, the development of a grounded theory does not require researchers to suspend or ignore existing theoretical knowledge, but encourages the development of grounded theories by drawing upon broad theoretical approaches beyond the same substantive area. Finally, grounded theory development “bridges the gap” between interpretive and positivist approaches. From an interpretive standpoint, data are collected qualitatively, sampling and analysis techniques are adopted, and the development of a theory is an inductive process. From a positivist standpoint, Strauss and Corbin provide systematic coding procedures, to which Goede and de Villiers suggest can be useful “to organize one’s data and to strengthen the scientific value of the emerging theory” [20], as well as deductive verification of findings through the constant comparison of data.

3.3. Methods Used

We describe our experiences with a well-known UK based computer games developer, VIS Entertainment Ltd. Over a period of six months, we observed the working methods and carried out unstructured interviews with the audio staff and other members of the development team. Extensive field notes were then coded using Ethnograph software. Visits took place once a week, depending on the availability of participants, and lasted approximately three to four hours at a time. Thus far, we have not posed structured questions; instead, questions were asked in an open fashion, to confirm observations, and were kept to a minimum to prevent disturbance. As the research progresses, we hope to conduct further interviews (structured and unstructured) and observations with other sound designers and their colleagues from similar games development companies to identify concepts and thus to develop a grounded theory.

4. IN THE FIELD

4.1. VIS Entertainment

VIS Entertainment Ltd was formed in the mid-1990s and quickly gained an international reputation. The company has around 120 employees working in offices in Dundee and Edinburgh in the UK, working on a variety of games for all the major platforms. Its greatest commercial success, State of Emergency, topped both the UK and US games charts simultaneously, selling over a million copies in the process. In May 2004, the company was acquired by Bam Entertainment Inc., a US-based developer, publisher and marketer of games software worldwide.

4.2. The Studio

Working in “the studio”, as it is referred to by members of staff we spoke to, the audio department is made up of three members: Steven, the Head of Audio, and two sound designers, James and Scott. They are responsible for the final audio production of current developments. This covers a wide spectrum, from recording and editing spoken dialogue through to the design of sound effects, ambiances and music. Steven will attend regular meetings with non-audio production staff to ensure he is aware of audio requirements, and will feed these expectations to James and Scott. Steven will also use these meetings to make his influence known, suggesting his own opinions for the use of audio.

Each sound designer will tend to concentrate on one particular development (at the time of writing, James was working on a game called Brave, while Scott was working on State of Emergency 2). The studio is based a short walking distance from the main office in Dundee, and is essentially first floor converted office space with stunning views over the River Tay and beyond. As only three employees work in the studio full time, it does appear at first glance to be slightly too large for their purposes. To compensate for this excess space, it appears to double as a storage facility for other offices – there is a great deal of shelving, some of which holds numerous issues of magazines such as Future Music, Mac Format and Studio Sound, all in cardboard folders and in order, as well as numerous manuals and boxes for various software and hardware. Larger, monitor sized boxes nestle behind the staff where there is space but, overall, the office is tidy and well maintained. All three staff share the views expressed by Marks,

---

Information is taken from the company’s website: http://www.visentertainment.com.

Names have been changed to protect the anonymity of participants.
who suggests sound designers “never have to wear a tie” [21]. Each of them will also spend a few days a week in the Edinburgh office, where they are able to carry out some day-to-day work and to discuss pressing issues with other (tie-less) members of the development team face-to-face.

Working for a larger and well established developer, the team has access to a wide range of contemporary audio tools. At the Dundee studio, two rooms are dedicated to recording; one is a dedicated vocal recording booth, while the other is a dedicated recording studio containing a plethora of audio and visual equipment including hardware synthesizers, sound modules, and various effects units alongside sundry items such as royalty free sample and TV theme CDs. Each sound designer has his own desk and storage space; Steven has a relatively small desk housing a desktop PC running Sonic Foundry Sound Forge 6.0 for sound editing, a digital mixing desk, a PlayStation 2 with a development kit attached, a television to monitor PlayStation games, and two audio monitors, not to mention copious notebooks and post-it notes.

4.3. The Sound Designer

Scott, a graduate in audio technology, is responsible for much of the sound design at VIS, and spends his working life composing sound effects, background ambience and recording/editing dialogue to ensure his designs “fit” into the particular game he is working on. In his words, he wishes to ensure the game “sounds like it looks like”. He suggests a fight scene as an example; as the action increases, he must ensure that this must be reflected aurally. However, abstract events require some additional thought. Pointing out a scene in which the main character is to blow up a safe, he rhetorically asks, “What sound should that be?” before demonstrating the resultant “smash” effect. Furthermore, he is conscious that he is restricted in what he can do; “you need to get the right sound to work, but it can be tricky for the programmers”. He can be creative, but must be aware that compromises must be made, particularly with just 2Mb of local sound memory available on the PlayStation 2.

Scott will normally be provided with a working (although not necessarily complete) version of the game known as the “build”. Essentially, the build is an interim functional version of a level or mission (levels are stages in the game the player must negotiate sequentially, while missions are “sub-levels” in which the player must complete a certain task within the level itself), and is a result of many months of highly collaborative work between game designers, artists (sound and vision) and programmers. Although constantly aware of the game design document (a document which Bethke describes as “the fun document that details all of the characters, the levels, the game mechanics, the views, the menus, and so on – in short, the game” [22]), Scott is constantly looking for opportunities to add sounds as he guides his character through the build. Background ambience will be added first; within a docks mission on State of Emergency, 2, for example, we hear the sound of a ship horn, squawking seagulls, and the splashing of the waves. This type of audio is purposefully designed to set the mood of the game, but also to linger almost unnoticed in the background. More general sound effects will then be incorporated into specific visual and non-visual elements in the game; the shooting of a gun, the sound of footsteps, the buzz of a helicopter not yet visible. The addition of dialogue, music and interface sounds (i.e. sounds which are not perceived as belonging to the diegetic part of the environment, e.g. changing controller settings, or sounds to indicate the loading or saving of games [23]) will ensure that the combination will contribute to the overall game playing experience.

So, how is this experience achieved? Let us illustrate using an example from the field study. The main character has to shoot at a fire extinguisher. Scott picks up his game controller and moves his character up to the fire extinguisher and begins to shoot – there is a visual effect, a plume of smoke, but no sound. Scott accesses Sound Ideas5, a commercially available sound library which contains thousands of sound effects from animal sounds through to traffic and weather effects. He types “Fire Extinguisher” into the search field. Several files are found, so he copies them across to Sound Forge, creating five tracks of appropriate sounds. The next few minutes are spent moving sounds around, playing them, cutting them, adding reverb, changing pitch. He moves his left hand over the PlayStation game controller and his right over the key to play the sample. He returns to the fire extinguisher and begins shooting, tapping the keyboard to ensure the sound plays in synchronization. He replays the scene over and over again, a process which takes over half an hour, adjusting the sample on Sound Forge. Once he is satisfied that his sound fits the visual action, he saves it in Audio Interchangeable File Format (AIFF) format and accesses a bespoke in-house piece of audio software which allows him to incorporate his sounds into the build “live”, whether the relevant platform is a PlayStation 2, Xbox or GameCube. The use of this application means that Scott need not personally access the underlying code at all; indeed, Scott mentions that his only programming experience was at university – “I wouldn’t be able to cope without this!” he admits. It is certainly arguable that implementation of sound within the build would take much longer without this software.

However, Scott does not spend his whole day on sound design. He tries to remain aware of problems faced by other members of the development team, and thus spends a great deal of his day communicating to other members of staff via e-mail, often spending large amounts of time on a single e-mail to get his point across, particularly if he feels changes are to affect what he does; “it’s amazing the amount of time I spend on e-mails”, he mentions during one visit. Similarly, if he finds that sound is not playing when it should be, he must negotiate with the programmer involved to find out why this is the case. Hence, the process of sound design is a highly co-operative activity. Like his colleagues, he has milestones to comply with, and thus Scott relies on other staff members, for example graphics artists, to complete their individual tasks before he can begin the process of design and implementation of audio. He relies on programmers to be able to fix problems so that he can ensure his designs are being used properly.

To illustrate, during one visit to the Edinburgh office Scott is asked to present his work-in-progress to Steven for his opinions. Scott runs through a level pointing out what he has done and where problems which he has been unable to fix still arise. Steven watches the action intently, pointing out situations in which sound is not present but could be added. For example, the player has to abseil down a building; at the moment, although this can be seen visually, there is no audio to compliment this task, and thus Steven asks if an audio effect can be added. Steven also points out a small problem. Why are there no screams when an enemy is shot? Scott replies that they

---

4 Fieldnote extracts are edited versions of notes taken in the field.

5 http://www.soundideas.com
are too far away from the main character, and thus dialogue there would be inappropriate. Steven disagrees – “there is a distinct sense of not a lot going on,” he replies, “we need to add vocal stuff”. In essence, Steven is highlighting a mismatch between the visual and the aural; visually, there is “a lot going on”, but this is not matched aurally. This echoes Michel Chion’s reference to sound as “added value”, i.e. “the expressive and informative value with which a sound enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression ‘naturally’ comes from what is seen, and is already contained in the image itself” [24].

This is not the only problem Steven picks up on. Both the main character and the other characters in the game use weapons which trigger the same machine gun sound effects, which have priority over other effects. The use of priorities allows for higher priority sounds which produce a greater audio effect, e.g. weapons or bombs, than lower priority sounds, e.g. footsteps. In this case, however, there is a problem. When several characters appear on the screen at the same time, the cacophony of machine gun sounds drown out the lower priority sound effects, which include vocal effects made by characters when they are shot. Steven suggests that the scene is adjusted so that either other characters use different weapons with less priority, for example pistols, or so machine gun effects are not played constantly and thus prevent this cacophony. Here game audio is driving other areas of game development; for example, if enemy characters were to use pistols instead of machine gun fire, the artwork would have to be changed to reflect this.

5. DISCUSSION AND FUTURE RESEARCH

The previous section provides an insight into the world of sound design during the course of the game development process. Are there lessons to be learnt for more general auditory interface design? To begin this discussion, let us examine some of the issues which arose during the field study thus far, and where there is potential for future research.

Firstly, communication makes up a considerable part of a sound designer’s day. Outside of game development, there is still an assumption that sound design and implementation is carried out towards the end of the development process. For example, Deutsch states: “The continuing weakness in the area is due to the unnecessary constraints in the production process, based upon old methods of operation. These constraints place sound near the end of the chain of production.” [25]. However, while this may be true of some developers, this is not the case at VIS. As has been seen, Scott communicates regularly with other members of the development team, and with Steven, to ensure that implementation problems can be prevented at an early stage. This is facilitated through the use of regular meetings, game documents, milestones, e-mails, and so on. Within auditory display research, it may be fruitful to investigate how the sound designer should fit into the overall design process of a user interface. How do we encourage increased collaboration between user interface designers and those working on audio? Could this lead to interfaces which best serve the user by using audio which has been developed iteratively rather than as an “add-on”? Participatory design [26], whereby the user is central to the project, often involved in setting design goals and developing prototypes, is one example of advancements made in people-centered design but, as Macaulay and Cerar [27] suggest, participatory design techniques such as paper prototyping and storyboarding are more suitable to GUI design rather than auditory display design. How could these techniques be adapted to help a wider range of people participate in the design of auditory interfaces?

Secondly, examining issues of annoyance may be fruitful; a major challenge for the sound designer is to ensure that the ambiance will remain interesting for the many hours a game may be played in a single setting [28]. While annoyance has been discussed previously within auditory interface research (e.g. [29]), it is arguable that there is further scope for investigation within computer game sound design to identify how sound designers ensure their designs are not repetitive over a long period of time. Understanding such issues may have implications towards how designers should approach auditory interfaces which are used over similar time periods.

The use of bespoke application software to integrate sounds within a game “live” may also interest designers of auditory interfaces. Through the use of this application, Scott is able to monitor his work immediately, and hence is able to identify discrepancies, contextual or otherwise, very quickly. While the lack of flexibility of various commercially available tools to integrate audio into user interfaces has already been discussed (e.g. [30]), it is arguable that the use of similar tools within auditory display design could facilitate a more iterative approach to sound design, as designers of auditory displays could monitor their work constantly during development rather than towards the end.

In our research thus far, we have concentrated on the use of sound effects and ambience within computer games. Of course, music plays a large part within contemporary computer games, and thus research into its use could be useful. The use of music within more general HCI has already been considered (e.g. [31]) and thus a logical next step would involve investigating its usage within computer games to identify possible approaches for more general interaction design.

6. CONCLUSION

The central claim of this research is that an understanding of approaches to sound design within contemporary computer game development can inform the ways in which designers approach contemporary auditory interfaces. We began by proposing the similarities between computer games and more general applications. Firstly, in both cases, sound is perceived as being in “third place” behind visual and interactive elements. Secondly, within computer games, appropriate sound is perceived to increase playability; within general interfaces, a successful auditory interface is one in which navigation is improved through the use of sound. However, there is one major difference which is pertinent: while it is conceivable to have a silent interface within more general applications, silent commercially available computer games are almost totally non-existent. Hence, an understanding of the use of sound in computer games can inform the designer with an understanding of how sound could be used within contemporary auditory interfaces.

An understanding of the approach to sound design within the overall computer game design process can inform the approach to auditory interface design. Through our field study, we discovered that sound design is highly collaborative, and plays a significant role within the development lifecycle. Whilst more research is required, it is necessary to treat sound within other forms of interface design in a similar fashion. It is necessary to consider how this approach should be achieved – we need to identify appropriate methods for the design of
interfaces in which sound design is a constantly evolving part of the entire development process.

7. ACKNOWLEDGEMENTS

The authors would like to thank all of our collaborators at VIS Entertainment for their time, cooperation and patience. This work has been funded by the UK Engineering and Physical Sciences Research Council (EPSRC).

8. REFERENCES