Beyond the Screen

What we can Learn about Game Design from Audio-Based Games

Karen Collins
The Games Institute
University of Waterloo
Waterloo, Canada

Bill Kapralos
Faculty of Business and Information Technology
University of Ontario Institute of Technology
Oshawa, Canada

***Abstract*— Sound plays a vital role in the communication of information in computer applications and entertainment. In video games, sound has many important functions, from communicating important information to the player to helping emotionally invest the player in a narrative. Despite its importance, sound has typically played a lesser role in interactive media, where emphasis has been placed on the visual scene and the generation of realistic graphics. Audio-based interactions in games have recently been explored as an alternative to visual-based media, and can offer us many insights into ways to better implement sound in audio-visual media. Here we provide a survey of existing audio-based games with a special focus on Beowulf and *Papa Sangre*. In the process of doing so, we describe approaches to changing the ocularcentrism of games in favour of a more balanced presentation of audio-visual information.**

Keywords-sound; audiogames; games; design; Papa Sangre

I. INTRODUCTION

As in the real world, sound plays a vital role in the communication of information in computer applications and entertainment, such as office application software and video games. In interactive applications such as virtual environments and simulations, auditory cues can help a user to orient themselves, add a better sense of presence or immersion, compensate for poor visual cues (graphics), and at the very least, add a pleasing quality to the simulation [1]. Sound in entertainment applications such as video games has many important functions: It can help to communicate important information to the player ("quick! Take out your weapon, there's an enemy around the corner!"); it can serve as a sound symbol or leitmotif; it can help to situate the player in a specific location; it creates a sense of realism. Sound effects associated with particular visual imagery (such as footsteps, a door opening, glass breaking, a ball bouncing, etc.) can lead to a greater sense of presence, realism and quality. Sound also plays an emotional role that helps to immerse the player in media, to emotionally invest them in a narrative.

Although sound is a critical element in perceiving our environment, it is often overlooked in interactive applications such as video games and virtual environments, where, historically, emphasis has been placed on visuals (graphics). Fortunately, there is a growing effort to ensure that interactive applications and digital media in general is accessible to those with disabilities of all types including visual impairment. The majority of the visually impaired rely on, and are well practiced with the use of hearing to gather information about their surroundings. This paper provides a survey of existing audio-based interactions in video games. In doing so, we describe approaches to changing the ocularcentrism of digital media in favour of a more balanced presentation of information. What can we learn from audio-based interactions that may improve our approach to developing audio-visual media? To answer this question, our focus here is strictly on audio-based interactions, rather than some of the audio-haptic multimodal interactions for the visually impaired that have recently been developed. We begin with an overview of audio-based games, and provide a detailed look at two games in particular, *Beowulf* and *Papa Sangre*. We conclude the paper with an overview of what audiovisual media developers and designers can take from these audiocentric experiments.

II. A BRIEF OVERVIEW OF AUDIO-BASED GAMES

Audio games includes audio-only computer games (that is, where there are no visuals) as well as audio-based games (where visuals are included but not the focus). The idea of non-visual interaction with a computer interface has been a popular endeavour for human computer interaction designers. Many early interfaces designed for the visually impaired initially used synthetic speech or a combination of musical tones and speech [2]. Similarly, audio augmentations of visual interfaces have existed since Bill Gaver's experiments with *SonicFinder* on the Macintosh computer in the late 1980s [3]. Although accessibility in user interfaces has been the subject of much work in interaction design, research and technology in other forms of media accessibility has lagged behind. This is in part due to the view of academics that entertainment is not as "important" as other software applications in terms of accessibility, and in part because the entertainment industry has not seen the financial value in incorporating accessible options into audio-visual media.

However, attitudes towards games are changing, and video games are now being applied and used in a variety of teaching and learning situations, replacing/complementing traditional teaching methods. Recently, educators have started to take notice of the rich learning context the current characteristically long, challenging, and complex console-based video games provide players. In addition to serious games and educational games that serve a direct function (learning, rehabilitation, amongst others), games play an important role in social play and youth culture, and the creation of accessible games can ensure social inclusivity and reduce feelings of isolation or

difference amongst those with impairment [4]. Despite a number of accessibility-related improvements particularly with respect to user interfaces, much work remains with respect to developing accessible video games. That being said, making video games accessible to the visually impaired poses greater challenges than software/user interface accessibility in general given its complexity. Furthermore, unlike basic user interfaces, emotions are an important part of video games, and have additional requirements beyond accessibility to information.

The idea of audio games dates back to at least *Valhalla and the Lord of Infinity*, a "speech adventure" game for the Amiga, although it still relied heavily on graphics (Vulcan Software 1994). A few years later, *Real Sound – Kaze no Regret* was released, a commercial audio game designed specifically for visually impaired children and released by Sega in 1999 for Dreamcast and Sega Saturn. The game was designed for the visually impaired, and included Braille-cards, although the release on the Dreamcast included an optional Visual Mode. Essentially, *Kaze no Regre*t was an interactive radio drama, similar to a choose-your-own-adventure style audio book. Other early audio games were similarly drawn from early text-based games as they were easy to convert to "audio-only" given they did not rely on visuals and at times did not include any visuals. Several audio-based or audio-only games have now been adapted from popular existing games. A brief survey of these games in the following section will highlight some of the unique abilities of audio games.

III. SPATIAL ARRANGEMENTS AND MENTAL MAPPING

Audio-based games have demonstrated the ability of players to use sound to create a mental map of the spatial environment of the game. Sánchez, Lumbreas and Cemuzzi [5] developed *AudioDoom*, an audio only game intended for young children based on the original *Doom* game (id 1993). *AudioDoom* is an interactive story that can be changed by the players' actions. In the game, the environment is arranged as a spatial connection of locations; some locations can include others, or be connected to others in a hierarchy (such as a neighbourhood containing houses, and each house containing rooms). Within each location, the objects, items, or other characters are presented in voxels (volume elements), and spatialized (to provide the player the impression of a sound at a particular location in three-dimensional space) using head-related transfer functions (HRTFs).[1] For example, the sound of a door can be played in the left most voxel, while a key to be picked up is on the right, and a monster can be moving between voxels. This is presented as the associated sound coming from a certain direction. The player interacts with the game by pointing, either with a specialized joystick or a regular keyboard, to the voxel of interest. Based on what is currently in that particular voxel, the action will be interpreted differently, such as opening the door, picking up a key, or shooting at a monster before it moves to a different voxel. A user study was conducted with seven visually impaired children to test how successful the sound was in conveying spatialized information. The children were asked to reconstruct the space, including orientation of rooms, doors and objects, using Lego-type building blocks. Most children exhibited a large degree of accuracy in representing the spatialized game space, showing that spatial information can be conveyed well through the use of spatialized sound.

An *Audio Platform Game*, based loosely on *Super Mario Bros*. similarly tested participants' abilities to map game levels [6]. Sighted players were assigned to either audio-only or audio-visual output, and visual impaired players were assigned to audio-only output games. It was discovered that it was entirely possible to be as effective in terms of creating mental-maps even for the sighted players playing audio-only versions of the game, suggesting that mental mapping may be an important unconscious process during any game playing [6]. With adaptations of games like audio versions of *Space Invaders* [7], it is clear that even two-dimensional visual games are better visualized with the incorporation of three-dimensional sound.

Three dimensional (spatial) sound allows a listener to perceive the position of sound sources, emanating from a static number of stationary loudspeakers or a pair of headphones, as coming from arbitrary locations in three dimensional space. Spatial sound technology goes far beyond traditional stereo and surround sound techniques by allowing a virtual sound source to have such attributes as left-right, back-forth and up-down [1]. The foundation of spatial sound rests on the ability to control the auditory signals arriving at the listener's ears such that these signals are perceptually equivalent to the signals the listener would receive in the environment being simulated [8].

In contrast to the visual sense, hearing is omni-directional, and one of the main benefits of audio games is the capability of a 360˚ field of interaction using spatial sound techniques. However, despite the benefits of spatial sound, its generation is not a trivial matter. This is particularly so when considering the reconstruction of a listening environment by physical or mathematical modeling means, taking into account the acoustics of the environment and the characteristics of the listener. Here, to recreate a sound in a three-dimensional space, sound source material must be filtered with the appropriate head-related transfer functions and room impulse response (or RIR, which captures the reflection properties, diffraction, refraction, sound attenuation and absorption properties of a particular room configuration (i.e., the "room acoustics") dynamically to account for listener and sound source movements, in real-time. Filtering is accomplished by a convolution operation. However, convolution is a computationally expensive operation especially when considering the long filters associated with HRTFs and RIRs (filters with 512 coefficients are not uncommon) thus limiting their use to non-real-time applications.

[1] The head related transfer function (HRTF) refers to the distance- and direction- dependent filtering effects of a sound reaching a listener due to the listener's head, shoulders, upper-torso, and most notably, the pinna of each ear. The HRTF corresponding to a particular location in three-dimensional space can be measured or modeled and used to filter a sound. When the filtered sound is presented to the listener, it provides them the impression of a sound originating from the HRTF measurement location.

Funding for this research was provided by the Social Sciences and Humanities Research Council of Canada.

Many spatial sound games allow the user to explore an imaginary three dimensional world of some form. *Sleuth: An Audio Experience* is another spatialized audio-based game [9], essentially an adaptation of the popular board game *Clue*. The player hears a short narrative providing them with information such as names of the 10 guests, and a list of weapons. The detective navigates the rooms attempting to listen and move towards key sounds (ambient and conversations). The authors discovered two primary difficulties with the game. The first was differentiating speakers in conversations: This could, of course, be rectified by using speakers with unusual voices or accents. The second was in identifying rooms with somewhat ambiguous ambiences. Some rooms were obviously defined, such as a kitchen with a leaky pipe, although the kitchen in this case also contained a television set, confusing the players. Having very conventional and stereotyped sound effects for ambience would, therefore, rectify this problem. The authors note that "A distinctive, salient ambience was also essential for identification of a given room. Subjects demonstrated this concern by consistently identifying the living, dining and billiard rooms, but they had more difficulty identifying the subtle or non-characteristic ambiences of the lounge, study, and kitchen" [9].

IV. LAIR OF BEOWULF

Nordlinder created the game *The Lair of Beowulf* with colleagues at the Interactive Institute, [10][2] based on their proposition that immersion, and particularly game immersion, is not necessarily about believable visual cues, but rather that it is very possible that visual representations in fact detract from the immersion. In other words, they believe that "less is more", and that creating situations that hint at something will immerse the user much more than actually showing it. In the game, the player takes on the role of the character Beowulf. Beowulf enters the cave system where Grendel (another game character) resides, and needs to navigate the cave system while avoiding various monsters, to find and slay Grendel. But upon entering the caves, Beowulf's torch is snuffed out by a gust of wind, and he must therefore navigate by hearing only. All that is presented on the game screen is a map indicating which part of the cave the player has already explored, and the rest is achieved using sound. The game is played from a first person audio perspective, where the player hears sounds as Beowulf would hear them. The sounds include Beowulf's footsteps, which change according to the surface he is walking on, the size of the area he is in (according to the reverberation patterns), and breathing sounds, which are useful for detecting walls (reverberation is cut down as the wall is approached). There are also ambient sounds, such as the sound of wind in a tunnel, or dripping water in a cave, sounds of living creatures, and combat sounds to provide feedback to the player.

User tests were conducted to measure the game's effectiveness. First, players were asked to navigate the game world and find seven specific caves within a reasonable amount of time in order to demonstrate that it is possible to navigate a game world based on spatialized sound only. Secondly, players were asked to describe the game world in order to gauge whether the game generated a rich mental image. Both tests were successful, showing both that players were able to navigate the world using sound (and a limited explored-area visual map), and the players came away with a very imaginative description of the cave system. The authors concluded that "When using sound and graphics the sound is often only a complement to the graphics and does not need to be descriptive, in contradiction to when using only sounds then it is important to have a descriptive sound if the intention is to make the sound fit to the game context" [10].

Beowulf's use of sound can be confusing for the player, and this is evident by the need to include a visualized map: without clear directional beacons in the soundscape, there are many parts of the cave system where the player is left trying to figure out which direction they are facing unless they use the visualized map. When there were clear positional markers—such as a waterfall—the player could use the location of the sound to determine their own location in the level. The problem with this, however, is that the waterfall sound is used more than once in the same level, so without a map, the player was left wondering if they'd back-tracked. Moreover, there are some sonic symbols used that are without a clear meaning at first: on the first level, the user may be unsure if the horn sound implies that they slayed a beast, or if the game was over and they won. Some of these issues have been overcome with the more recent but similar game, *Papa Sangre*.

V. PAPA SANGRE

Papa Sangre is a popular iPod/iPad audio game produced by Somethin' Else (2010), similarly relying on binaural audio technology to create a spatial environment in headphones. The player navigates through five palaces in the "land of the dead", without the use of their eyes, by tapping on feet on-screen. The game's story is told through sound, with each castle having its own sonic identity, and different forms of monsters represented through sound effects. The main goal of the game is survival, as the player wanders the castles collecting musical notes that commonly hide behind one of the many roaming monsters. Monsters respond to player-generated sound, so if the player moves too quickly, or steps on an object that makes sound, they will be chased and killed. The player controls the pacing, and since walking too loudly can trigger a demon to recognize and chase you, the player is in control of the arrival of monsters. Moving too quickly can cause the player to trip, leaving the player open to the monsters. The player, then, must have patience and move slowly through the space. This is not a game that can be rushed, and by controlling the pacing in this manner, the game forces the player to listen while they move, to pay attention to their own sounding bodies in the virtual world. It is not only the sounds other characters or objects are making, but the player's own sounds to which they must actively listen.

Papa Sangre overcomes some of the problems that were apparent in *Beowulf*. It has a narrated helper that explains each level before the player enters the room, orienting the player in the space (e.g. "Collect the trail of musical notes, but a hog is patrolling the room from left to right in front of you."). Each room has its own sonic beacons, the musical notes that the

[2] We played the *Lair of Beowulf Light* version available on the Interactive Institute website: http://www.tii.se/projects/beowulf.

Figure 1. *Papa Sangre* orientation wheel.

player must collect, but it also has other locational markers: sleeping beasts and so on. By using these sonic markers, a map is not necessary for the player: the player can always orient themselves by the continued sounds. By spinning the 360˚ orientation wheel (Figure 1), the player can turn around smoothly, re-orienting themselves towards the sound by lining up the amplitude in both ears equally.

Both Papa Sangre and Beowulf were able to cleverly limit the amount of sounds present while still maintaining a sense of the environment. Too many sounds can be detrimental to usability and enjoyment, and can lead to confusion. It is likely that coherence will reduce this effect, but there has been little research into how many sounds we can reasonably process under differing contextual circumstances. Sound designer Walter Murch refers to the moments of too many simultaneous sounds as a "logjam" and asks how to choose which sounds should predominate when they can't all be included? In his mix of *Apocalypse Now*, he found that six simultaneous sounds in a mix were the maximum, but he also notes that when sounds are in harmony with each other (e.g., in the same harmonic spectrum), this can change how we perceive the number of sounds, and can change this effect [11]. *Papa Sangre*'s sound designer Nick Ryan found that the most that they could get away with was three: "We were limited by audio cognition – our ability to perceive the directionality of a sound and space. To imagine ourselves to be in an audio environment we can only understand a number of sound sources at one time – we had to cut that down to three otherwise it was too confusing" [12].

VI. Conclusions: What can we learn about Game Design from Audio Games?

We have presented a summary of some of the existing audio-based games. These games have taken a variety of approaches, but from each we can learn something about the roles that sound plays in human-computer interaction, and we can then implement these findings back into audio-visual media to create an enriched experience, facilitate more efficient and effective interaction, and increase enjoyment. We can summarize these findings and their implications for game and interface designers as follows:

1. Orientation and spatial awareness: One of the advantages of game sound is its ability to extend the game beyond the screen. While the player may have a limited window of a 360 degree game to view at any one time (in first person shooters, this is usually between 65 -80˚ [13]), they can sonically hear the entire 360 degrees. We have seen that game players were able to create visual mental maps of space based solely on hearing spatialized sound. This capacity of sound to create a visual space in our minds is perhaps one of sound's most untapped abilities. Using only sound, we can create a space that expands beyond our small mobile displays or television screens. This mental space created by sound can also be more powerful than the audio-visual space created for us. However, the ability of players to orient themselves in this un-seen space is dependent on how the sound design is implemented. Players need a distinct auditory beacon towards which they can head. This will also help them to create a mental map of the entire game space, ensuring greater efficiency in their own movements.

2. Creating consistency: Another important element relating to the perceptual properties of sound is the context in which the sounds were heard. Summarized by Friberg and Gärdenfors, "The sound objects of an audio game are always heard in a context. Firstly, there is the context of the game plot, where the individual sounds should be intelligible and distinguishable. Secondly, there is a musical context where each sound object is accompanied by the other sounds of the game. The later context is seldom emphasized in game audio design, especially not in games accessible to visually impaired players" [14]. Sounds work best when they are part of a coherent whole—whether notes of a scale or raga, or part of some other coherent conceptual environment. For game designers this task can prove difficult. Many game development companies have teams of people working on sound, some of whom may be outside contractors and have little contact with each other. The sound designers, for instance, may not work closely with the composers. In terms of the suitability of sounds and the creation of a sense of coherence, therefore, it is important that everybody contributing to the sound of a game be in contact and working together to create the overall soundscape.

3. Less is more: Sound and image can have an amplifying effect on each other. But even on its own, sound can have a powerful ability to create mental imagery. As is often found by people who think that "the book was better than the movie", the imagination can be far more powerful than having everything shown to us. The lack of images is one of the elements of the game that makes it more frightening than most games, as players comment: "And even though you can't see anything, or maybe because of it, Papa Sangre is terrifying."[3] "In fact, every time you hear anything in Papa Sangre your heart races, even when it shouldn't. Babies crying, telephones ringing – it's all scary to me now." [4] Too

[3] http://www.gamezebo.com/games/papa-sangre/review

[4] http://www.148apps.com/reviews/papa-sangre-review/

often, games rely on fancy graphics to portray detailed information that could more effectively be portrayed through sound. In particular, emotionally important information is probably more effectively conveyed through sound.

4. Focus on the player's character: What *Papa Sangre* has accomplished with sound is to elevate its importance in gameplay: not only is this game not playable with the sound off, but the player must constantly actively listen while they play, to both their own sounds and the sounds in the environment. This highlights one of the elements that is missing from many video games: our own avataristic self-awareness in the game. As players, we rarely pay attention to our own character in first-person perspective: we may speak through their body or see our hands on-screen, but we don't often focus on how fast we are breathing or listen to our own footsteps. Stealth games like *Thief: The Dark Project* (Eidos, 1998) have used sound cues helped to inform the player of nearby enemies, and also helped to inform nearby enemies of the player. Guards usually alerted the player to their presence through sound by walking heavily, singing, or whistling. If the player, playing as thief Garrett, made too much noise, nonplaying characters would be alerted to his presence. The player had to walk softly, muffle footsteps with moss and stay away from hard pavement, grates, or tiled floors. If a guard or other enemy was alerted, the player had to remain still and hide until the enemy gave up searching for him. Such use of sound is still significantly under-used in games, and could offer the player another interesting dimension of play to focus on, and help them to better identify with and role-play their character.

In summary, we have demonstrated that we can learn much from creating and studying audio-based games. We can incorporate these lessons into video game design to enhance the role that sound plays. We have also shown, however, that there are many areas of sound research that remain unexplored or under-researched. These areas offer many dimensions of future research possibilities that can help us to understand not just sound in isolation, but the role that sound plays in audio-visual media today.

REFERENCES

[1] M. Cohen and E. Wenzel. "The design of multidimensional sound interfaces." In W. Bareld and T. Furness, editors, Virtual Environments and Advanced Interface Design, chapter 8, pp. 291–346. Oxford University Press Inc., New York, NY. USA, 1995.

[2] M. Cooper and H. Petrie. "The design of auditory interfaces for visually disabled users." In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '88), pages 83–88, Washington, DC. USA, May 15-19 1988.

[3] B. Gaver. "The SonicFinder: An interface that uses auditory icons" Human Computer Interactions 4(1), pp.67–94, 1989.

[4] D. Archambault, T. Gaudy, and K. Miesenberger. "Computer games and visually impaired people." UPGRADE, The European Journal for the Informatics Professional, 8(2), pp.43–53, 2007.

[5] J. Sánchez, M. Lumbreras, and L. Cernuzzi. "Interactive virtual acoustic environments for blind children: Computing, usability, and cognition." In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 65–66, Seattle, WA.USA, March 31 - April 5 2001.

[6] M. A. Oren, C. Harding, and T. Bonebright. "Evaluation of spatial abilities within a 2D auditory platform game." In Proceedings of the 10th International ACM Conference on Computers and accessibility (SIGACCESS '08), pp. 235–236, Halifax, Nova Scotia, Canada, October 13-15, 2008.

[7] R. J. McCrindle and D. Symons. "Audio Space Invaders." In Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies, pp. 24–31, Alghero, Italy, 2000.

[8] D. B. Ward and G. W. Elko. "A new robust system for 3D audio using loudspeakers." In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2000), volume 2, pp. 781–784, 2000.

[9] T. M. Drewes, E. D. Mynatt, and M. Gandy. "Sleuth: An audio experience." In Proceedings of the 2nd International Conference on Auditory Display (ICAD 2000), Atlanta, GA. USA, April 2-5 2000.

[10] M. Nordlinder. "Lair of Beowulf: A study of 3D positional sound in an audio mostly game." Master's thesis, Department of Computer Science/ Media Technology, Lulea University of Technology, Lulea, Sweden, 2007.

[11] W. Murch. "Dense clarity - clear density." Transom Review, http://transom.org/?page_id=7006.

[12] "Papa Sangre: The sonic iPhone Horror game you've been looking for". The Guardian PDA: The Digital Content Blog. December 20, 2010. http://www.guardian.co.uk/technology/pda/2010/dec/20/papa-sangre-game-audio

[13] Stevens, R. and D. Raybould. The game audio tutorial: A practical guide to sound and music for interactive games. London: Focal Press, 2011.

[14] J. Friberg and D. Gärdenfors. "Audio games: New perspectives on game audio." In Proceedings of the ACM International Conference on Advances in Computer Entertainment Technology (ACE '04), pages 148–154, Jumanji, Singapore, 2004.