

**INTERACTIVE AND ADAPTIVE AUDIO  
FOR  
HOME VIDEO GAME CONSOLES**

by

**Jonathan B. Fish  
M.A., Simon Fraser University, 2003**

**THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF ARTS (COMMUNICATION)**

in the School  
of  
Communication

© Jonathan B. Fish 2003  
SIMON FRASER UNIVERSITY  
July 2003

All rights reserved. This work may not be  
reproduced in whole or in part, by photocopy  
or other means, without permission of the author.

**APPROVAL**

**NAME:** Jonathan Fish

**DEGREE:** MA

**TITLE OF THESIS:** Interactive and Adaptive Audio for Home Video Game Consoles

**EXAMINING COMMITTEE:**

**CHAIR:** Prof. Gary McCarron

---

Prof. Barry Truax  
Senior Supervisor  
Professor, School of Communication, SFU

---

Prof. Stephen Kline  
Supervisor  
Professor, School of Communication, SFU

---

Prof. Kenneth Newby,  
Examiner  
Assistant Professor, Interactive Arts, Computing Arts and  
Design Sciences, SFU Surrey

**Date:**

July 28/03

## PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

### Title of Thesis/Project/Extended Essay

Interactive and Adaptive Audio for Home Video Game Consoles

---

---

---

---

Author: \_\_\_\_\_  
(Signature)

\_\_\_\_\_  
(name)

July 28 / 03  
(date)

## ABSTRACT

The video game industry, more suitably referred to as interactive entertainment, serves to immerse players into a fantasy universe for an engaging, virtual, participatory experience. While visuals have been the focus of previous studies of the video game medium, recent advances in video game audio demonstrate that game sound is no longer merely an underlying component in crafting the overall, interactive experience.

The thesis analyzes the interactive nature of video game audio, with its real-time audio tools that allow audio components to adapt and conform to random gameplay actions taken by the individual player. Specific games from the different genres of First-Person Shooter, Sports Simulator, and Fantasy Sports games are studied using soundscape study concepts to determine how interactive and adaptive audio functions spanning different interactive game audio constructs. The methodology draws its data from multiple sources, including personal observations, literature surveys, case studies and interviews with professional video game sound designers. The dissertation brings together the discrete areas of soundscape studies (including acoustic and psychoacoustic concepts), video-game audio technology, and theories of interactivity in order to lay the groundwork for the audio-textual analyses of the games examined.

The thesis argues that interactive audio, within the immersive context of the virtual gaming environment, creates a feedback cycle which serves to empower the participant, and that the surrogate interactive soundscapes of video game environments become vehicles that make the interactive partnership between user and machine possible. The study shows that a communicational approach in soundscape studies can be readily applied to video-game soundscapes similar to its use in deconstructing sonic environments in the real world. The role of audio interactivity in immersing the player is analyzed by isolating the sonic elements within the virtual sonic environments inherent in specific games, analyzing their functional role within the gaming context, and documenting the experiences of professional sound designers obtained during interviews.

Dedicated to my father,  
Dr. William W. Fish

## ACKNOWLEDGEMENTS

I would like to thank all those whose invaluable support made this thesis possible. Without the help of those listed, this labour of love could not have come to completion.

First and foremost, I wish to extend my sincerest gratitude to my mentor, Barry Truax, for his ongoing support, scrupulous supervision, expert advice, and insightful knowledge. I have been privileged to work with Barry, who has been a meticulous editor, but also whose patience and friendly support will never be forgotten. A special thanks to David Murphy, whose technical assistance (as well as moral support) served to make the presentation of this thesis possible. Thanks to Aaron Marks for his support in granting me permission to incorporate from his inspirational and invaluable text, The Complete Guide to Game Audio. I wish to express my gratitude to Kenneth Newby for his feedback and insight, and for a thought-provoking discussion. Thanks to Stephen Kline for his guidance prior to the thesis in helping me explore the area of interactive entertainment media, and the issues and debates surrounding the vibrant medium of video games. I would also like to acknowledge the friendly faculty at the School of Communication, as well as the rest of the helpful SFU staff for their services. I wish to express my sincerest appreciation to the sound artists at Electronic Arts Inc., for sharing their experiences and expertise, and for giving so selflessly of their time. I wish to also recognize Electronic Arts Canada headquarters in Burnaby for approving the NDA, making the interviews in this study possible.

Finally, I would like to show appreciation to my friends and colleagues who have supported me along the way, especially my beautiful, loving girlfriend, Joanne, for helping me find my motivation when suffering from writer's block, or just being there as a shoulder to cry on. Thanks to my wonderful family, for helping me bear the weight of this undertaking. Last but not least, thank you Dad, for being my inspiration and for always believing in me. I couldn't have done this without you, and I will always remember your strength and dedication as the driving forces motivating me to persevere. We did it, Dad!

# TABLE OF CONTENTS

Approval .....	ii
Abstract .....	iii
Dedication .....	iv
Acknowledgments .....	v
Table of Contents .....	vi
<b>Chapter I. An Interactive Medium .....</b>	<b>1</b>
1.1 Video Games .....	5
1.1.1 Video Game Audio: “We’ve Heard It All Before... Haven’t We?” .....	5
1.1.2 Video Game Audio: An Aural History .....	6
1.1.3 At Present: “Hear We Are” .....	12
1.2 Interactivity & Immersion .....	13
<b>Chapter II. Sound: An Ear-Opening Experience .....</b>	<b>23</b>
2.1 A "Sound" Approach .....	23
2.2 Speech - Music - Soundscape .....	27
2.2.1 Speech: More than Just Hearsay .....	27
2.2.2 Music: 'Organized Sound' .....	28
2.2.3 Soundscape: Keeping an Ear to the Ground .....	30
2.3 Psychoacoustics & Auditory Perception: 'Clairaudience' .....	34
<b>Chapter III. Interactive Audio in Video Games: More than Meets the Ear .....</b>	<b>40</b>
3.1 Video Game Audio: Genres .....	40
3.1.1 First Person Shooter Games .....	41
3.1.2 Graphic Adventure Games .....	42
3.1.3 Role-Playing Games .....	43
3.1.4 Sports Games .....	44
3.1.5 Music Simulation Games .....	45
3.2 Video Game Audio: Home Console Systems .....	46
3.2.1 Sony PlayStation 2 .....	47
3.2.2 Nintendo GameCube .....	48
3.2.3 Microsoft Xbox .....	49
3.3 Technological Advances in Video Game Audio: Loud & Clear .....	50
3.3.1 Linear Audio .....	50
3.3.2 Interactive & Adaptive Audio .....	56

<b>Chapter IV. Principles of Design: Hearing Is Believing.....</b>	<b>66</b>
4.1 Interacting with Interactive Sound Artists.....	69
4.2 Achieving Audio Interactivity .....	84
4.3 Applying Interactive Design Principles .....	89
<b>Chapter V. Case Study: Video Game Textual Analyses .....</b>	<b>92</b>
5.1 Methodological Framework .....	92
5.1.1 The Method: Playing By Ear .....	92
5.1.2 Chosen Games, Genres and their Significance.....	94
5.2 Game 1: SOCOM: U.S. Navy SEALs.....	94
5.2.1 SOCOM: Speech, Music & SFX.....	96
5.2.2 SOCOM: The Virtual Sonic Environment .....	98
5.2.3 SOCOM: Interactive Audio .....	101
5.3 Game 2: SSX Tricky.....	105
5.3.1 SSX Tricky: Speech, Music & SFX .....	106
5.3.2 SSX Tricky: The Virtual Sonic Environment .....	109
5.3.3 SSX Tricky: Interactive Audio.....	113
5.4 Game 3: MVP Baseball 2003 .....	118
5.4.1 MVP Baseball 2003: Speech, Music & SFX .....	118
5.4.2 MVP Baseball 2003: The Virtual Sonic Environment.....	121
5.4.3 MVP Baseball 2003: Interactive Audio.....	126
5.5 Conclusions .....	130
<b>Chapter VI. Interactive Audio in Video Games: An Ear-Opening Experience .....</b>	<b>133</b>
6.1 Conclusions & Impressions: We're All Ears .....	134
6.2 Issues Arising from Interactive Audio Design .....	137
6.3 New Relationships & Directions: Now Hear This.....	140
6.4 Audio Interactivity in Video Games: A Bright-Sounding Future.....	145
<b>Appendix A.....</b>	<b>149</b>
<b>Appendix B.....</b>	<b>162</b>
<b>Appendix C.....</b>	<b>167</b>
<b>Appendix D.....</b>	<b>171</b>
<b>Bibliography .....</b>	<b>174</b>



*It is frustrating to see so much attention being given to eye candy when sound has the potential to provide the gamer just the same amount of sensory cues. The sound... is an incredible adrenaline pumping experience that games need to deliver to the gamer.<sup>1</sup>*

-Greg Hill  
Sound Designer for Soundwave Concepts

## 1. An Interactive Medium

Video game audio, as is the case with sound in general when contrasted with visuals in society, has always taken a back seat to onscreen graphics. As with film and television media, sound and music are frequently the forgotten fundamentals in the design of video games. This comes as no surprise, as the sound in most forms of media entertainment becomes the sonic backdrop over which the visual elements are perceived. Even the name “video game” implies that they are a visual activity.<sup>2</sup> One of the reasons for this may be that sound affects the player more subtly than flashy visuals, and the auditory components are perceived on a more subconscious level. Much of this relies on an effective, but understated sound design to support the visuals, as usually the mark of a good sound design is that you don’t pay conscious attention to it, and cannot single it out in the overall production of the game. However, sophisticated advances in video game audio clearly demonstrate that game sound is no longer merely an underlying component. Heightened tension and constant shifts in mood, prompted by the audio component within current game titles, draw the gamer into the world created by the game designers, and serve to immerse the player into a virtual fantasy realm. Only more recently has sound design been recognized for its merits, and a big reason for this, as put forth by Keith Arem, game composer and sound designer for PCB Productions, has been the lack of home game system fidelity.

PC speakers and console systems had limited audio fidelity and kept the resolution of sound to a minimum. In recent years there has been a strong effort to enhance sound for games, and... have made great strides to address memory and bandwidth for audio. Dolby encoding, DVD drives, surround sound speakers, and increased memory have given sound professionals a new field to play on.<sup>3</sup>

It becomes clear to anyone who has played a video game (through a television monitor containing even remotely functional speakers) in the past few years, that those annoying, ear-piercing,

---

<sup>1</sup> Hill, Greg. "Capturing Engine Sounds for Games." Gamasutra, 2002.  
[http://www.gamasutra.com/features/20021030/hill\\_01.htm](http://www.gamasutra.com/features/20021030/hill_01.htm).

<sup>2</sup> Andresen, Gavin. *Playing by Ear...*, 2002. [www.gamasutra.com/resource\\_guide/20020520/andresen\\_01.htm](http://www.gamasutra.com/resource_guide/20020520/andresen_01.htm)

electronic bleeps and buzzes that once accompanied video game play are a thing of the past. One only needs to briefly listen to the rich, emotionally-charged, orchestral tapestry and vividly recreated, heart-pounding, adrenalizing militant soundscape from games like *Medal of Honor: Frontline* to truly appreciate an example of the advances being made by sound artists within this new, and deeply engaging interactive medium. Although sound effects and musical scoring in games are still underlying components that are taken for granted by most, it is obvious even to the most uncultured ear that there has been an unmistakable and dramatic shift in the audio component of home video games. The visual aspect of video games are well documented, as well as the striking expansion in graphical technologies,<sup>4</sup> but very little has been written to explain the advances in the auditory realm of this ground-breaking medium, that not only serves as a means to pass the time, but allows the participant to participate in a fantasy universe, ever-ready to immerse and engage the player. Even the slogan for Sony's *Playstation 2* home console game system, appropriately named the 'Emotion Engine', encourages the player to escape into the fantasy world created by the gaming industry, urging the gamer: "Live in your world. Play in ours.™"<sup>5</sup>

The subject area examining levels of engagement with interactive forms of media, namely video games and the Internet, is of importance at the onset of the 21<sup>st</sup> century. Long past are the days of Frankfurt School theories, like those of Adorno and Horkheimer, that argued consumption of media can only exist on a passive level as a method for controlling media consumers, and does not allow for active response or engagement with media. Interactivity is at the forefront of media implementation, and the medium of video games is leading the way, with new and innovative technologies to bridge the gap between user and machine, making video game technologies adaptive and reactive to gameplay initiated by the individual gamer. However, amidst the glory of interactive technology there lies a growing fear of video games and their power to engage and immerse avid players. A 'New Media Panic' (Kline, 2000) continues to build in society, as concerned people contemplate the possible 'negative effects' of interactive video games on regular game players. As with any stimulating practice, overindulgence can prove to be unhealthy, and longitudinal studies attempt to determine how much exposure to video games is too much. While it has been found that video games can serve as a powerful tool to educate and teach valuable skills, it is often compared to an addictive drug that creates a dependency and warps the minds of susceptible players who immerse themselves too deeply into

---

<sup>3</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p. xvii

<sup>4</sup> Special Interest Group Graphics (SIGGRAPH), 2003. <http://www.siggraph.org/gen-info>.

<sup>5</sup> Playstation.com. <http://www.us.playstation.com/games/SCUS-97134.asp>. © 2003 Sony Computer Entertainment America Inc.

the virtual world created by the gaming industry (Kubey, 1996; Selnow, 1984). The purpose of this thesis is to contribute to a growing body of knowledge and literature about video game technology, its evolution, and the issues emerging from this revolutionary, interactive medium. Various articles have pinpointed the visual components of video games, marveling at the evolution of graphics.<sup>6</sup> Consequently, previous research has provoked debates surrounding the potential negative implications of visuals projected on the screen. However, little to no research has been devoted to the auditory component of video games, even though interactive audio technologies have advanced or even surpassed graphical developments in recent years. In fact, I think it is safe to suggest that there are deficiencies in the existing literature about video game technology, as most do not adequately, or even remotely, address the issues pertaining to audio.

The purpose of this dissertation is to explore, observe, educate and inform about interactive audio advances in video game technology, and to examine specific games from different genres using soundscape study parameters to perform textual analyses in order to determine how interactive and adaptive audio functions. The methodology I have used draws its data from multiple sources. I have collected data by means of triangulating observations, literary surveys, and interviews with accomplished video game sound designers. Having examined pertinent aspects of the subject matter from various angles, I feel I have sufficient understanding of emerging protocols and audio algorithms to contribute to this fascinating, emerging area of study. My goals in this analysis were to gather information through data collection, and to broaden my knowledge of sound and its function within an interactive gaming context. The task was to bring the areas of soundscape studies, video game technology, and interactivity together in order to lay the groundwork for the case studies and audio-textual analyses to follow. Because of the ongoing advances in home video game audio technology, I found myself having to update the current information to correspond with innovative technological trends in the industry.

Developing the auditory model for the video game analyses also proved to be challenging in combining these previously separate areas of study, for which there was no existing template. It was difficult to limit myself to certain aspects, as I became increasingly immersed in applying theories of soundscape studies and interactivity to multiple video games, spanning various genres. Also, it was a test for me to refine my scope to integrate only statements from game designers that were directly pertinent to my study, as there was so much material I found to be fascinating. I also encountered limitations, in that I had to respect various ethical guidelines and non-disclosure

---

<sup>6</sup> These include: Gamasutra.com, *Features, Audio & Graphics* (various 1997-2003): [www.gamasutra.com](http://www.gamasutra.com); Graphicsworld.net <http://graphicsworld.net/gaming.htm>; Game Development Search Engine (GDSE): <http://www.gdse.com/servlet/gdse.links?s=65>; GameDevelopers.net <http://www.gamedev.net/>; Special Interest Group Graphics (SIGGRAPH) <http://siggraph.org/>, etc.

agreements (NDA), so as not to violate or jeopardize the reputations of the participants in the study and the companies they represent. I also had to severely limit my discussion of ethical issues arising with respect to young audiences because of stringent regulations with respect to ethical research practices and the problematic nature of informed consent from minors. These obstacles led me to my current study that contributes to a new body of knowledge on current advances and designs in interactive video game audio, as constructing a virtual, surrogate audio environment are designed to involve the player as a participant rather than a passive observer. In this thesis I will argue that the interactive construct of video game audio, through innovative technological trends, serves as a driving force on a parallel with visual graphics to propel the active human participant into an environment through which the gamer has influence of his or her own virtual destiny in the simulated, electroacoustic environment. Having documented and deconstructed the analytical categories for analysis in this study, and applying the theories of interactivity and soundscape studies to the current technological state of home video game consoles, I can then hypothesize about the nature of the interactive video gaming environment with respect to audio. It will become clear to the reader, by applying grounded theories from the diverse areas being brought together in this study, that interactive audio within the immersing context of the virtual gaming environment creates a feedback cycle which serves to empower the participant. I suggest that the surrogate interactive audio soundscape of the video game becomes the vehicle that makes this interactive partnership between user and machine possible. This qualitative undertaking is meant to be exploratory and understanding-oriented, for the purpose of highlighting the experiences of the participants, as well as articulating personal experiences. Having conducted textual analyses, having surveyed the literature, and having conducted interviews with video game audio experts to pinpoint emerging protocols, I will then further develop my own meanings, hypotheses and findings in the final stages of the thesis that have evolved and emerged throughout my immersion with this captivating topic.

Before examining interactive audio in specific game analyses, it is first important to acquire a better understanding of the interactive video game medium itself. By getting a sense of the rapid growth in video game audio technology, leading to the enormous aural sophistication and expansion of the games, consoles, and evolving video game industry itself, we will gain a renewed appreciation for the auditory component of this engaging and deeply immersive game medium. First, let us look at the evolution of video game audio since its conception over three decades ago.

## 1.1 Video Games

Over 30 years after the invention of the first video game, the electronic gaming industry has emerged as the fastest growing form of entertainment, propelling it into a multi-billion dollar-a-year business. The entire electronic gaming industry, both hardware and software, accumulated an estimated \$31 billion US last year,<sup>7</sup> making it the most profitable industry in the entertainment trade today. Home video game sales rose 21% in the U.S. in 2002 as indicated by NPD Group Inc., an expansion eight times faster than the projected growth of the American economy.<sup>8</sup> Video game technology has evolved at such a rapid pace that games that were state-of-the-art five years ago now seem primitive and laughable to today's avid gamer. Because the industry has become such a cash cow, the biggest names in the industry now have an abundance of resources at their disposal for expanding their technology, especially with colossal companies like Microsoft breaking into the video game console market. Much like the computer industry, what was groundbreaking the previous year quickly becomes obsolete the following year, as new technological innovations make way for the newest, sophisticated game trends.

### 1.1.1 Video Game Audio: "We've Heard It All Before...Haven't We?"

It is only in the past decade, an extremely dynamic period in the game industry, that the audio component of video games has become recognized as a key element in the design of modern video games. Audio is only now being considered by game producers and programmers to be an integral element in the overall success and impact of a given game, perhaps even equivalent to the visuals (although this may be a stretch for some). Audio designers are now being integrated as members of the designing and programming team from the early stages of game design, and game soundtracks benefit greatly from the constant overview and regular input of the sound team. Now, with the advent of adaptive audio technology, interactive composing and authoring tools, Dolby 5.1 digital surround sound, orchestral video game musical scores, and real-time online gaming (to name a few), there is no question that video game audio is finally garnering some much deserved attention. As I will discuss in the third chapter of this thesis, interactive, adaptive audio technology, in contrast to the previous sample-based, predetermined soundtrack model, now allows for unpredictable game audio that actually responds to particular "gamestates", and actions taken by the individual player. Before examining the current

---

<sup>7</sup> McCullough, Michael. Technology: "Entertainment goes Wireless", *Vancouver Sun*, Saturday, March 22, 2003. Business section, E1.

<sup>8</sup> Bloomberg, "Film-based games...", *The Vancouver Sun*, January 31, 2003.

technology, I feel it might be best to first look at how video game audio technology has evolved since its conception.

### 1.1.2 Video Game Audio: An Aural History

Before the early 1970s, there was little to no demand for video game systems. Only a select few had the opportunity to try computer games, as they were only available on huge, multimillion-dollar, room-sized mainframe computers. In 1953, scientists at Bell Labs had devised a brand new electronic device called the transistor, which led to the invention of the integrated circuit (IC), which are the basis of every computer of any size today.<sup>9</sup> Affordable games for home use would not have been possible without these early developments.

There are different accounts of who actually invented the first computer game. The consensus is that the first computer-game design was conceptualized (with the help of technical specialist Robert V. Dvorak) by William Higinbotham, head of Brookhaven National Laboratory's Instrumentation division, in 1958.<sup>10</sup> The game was appropriately named "*Tennis for Two*" (which featured a ball bouncing off 2 movable paddles at opposite ends of a tiny screen), and the console consisted of a transistor, and an oscilloscope with a 5" screen, perched on a black box.<sup>11</sup> Ralph Baer, a senior engineer at Loral Electronics in New York in 1951, built the first video game prototype during the summer of 1966.<sup>12</sup> His "Chasing Game" consisted of 2 square-shaped objects that could be moved on a TV screen. Baer, considered the Thomas Edison of the video game, developed this idea into the Ping-Pong game concept along with fellow technical engineers Bill Rush and Bill Harrison in November of 1967. In 1970, Baer demonstrated his idea to Magnavox, who signed an exclusive, licensed agreement to produce his "Brown Box". In March of 1972, Magnavox introduced the first video game console system, "*The Odyssey*", which used no memory or microprocessor. It was purely analog. Between August and December 1972, 100,000 systems were sold.<sup>13</sup> Up to this point, video games had absolutely no sound component.

Nolan Bushnell, who later became president of Atari Inc., had the opportunity to sample the Odyssey Ping-Pong game. He had a vision of the game being played by the masses. Nolan Bushnell founded the Atari corporation, along with UC Berkeley Electrical Engineering and Computer Science graduate, Allan Alcorn (now Senior VP and Chief Technology Officer of

---

<sup>9</sup> Ahl, David H. "Editorial." 1983. <http://cvmm.vg-network.com/vag1.htm>

<sup>10</sup> Golberg, Marty. 1996. "The History of Computer Gaming Part 1: Forgotten Origins." <http://www.classicgaming.com/features/articles/computergaminghistory>

<sup>11</sup> Research and Development of the US Dept of Energy, 1981. <http://www.osti.gov/accomplishments/videogame.html#pagetop>.

<sup>12</sup> Winter, David. 1996. "Pong-Story." <http://www.pong-story.com/rhbaer.htm>.

<sup>13</sup> Ibid.

*Silicon Gaming*), who designed the first game under the Atari name. The game was entitled *PONG*. The name was actually chosen specifically for the audio component of the game, which featured a hollow, ringing “bleep” sound each time the ball made contact with the paddles.

Atari took over from Magnavox as the leading home console distributor in 1974, and the PONG game was redesigned and sold under its Tele-Games label for Christmas in 1976. The new Atari Video Computer System (VCS), released in 1977 along with its nine compatible game cartridges, was said to have a chip that allowed for digital on-screen scoring and “attractive sound.”<sup>14</sup> The “attractive” sounds being referred to were fixed, invariable, harsh, boring sonar “bleep” noises that pale in comparison to today’s lowest standard. Although the sounds consisted of primitive, hypnotic, computerized blips, they will always be revered as innovative, revolutionary, and the foundation for the complex video game soundtracks we know today.

From 1977 onward, game sound began to evolve gradually from year to year. It is important to note that early video game designers had a low budget to work with, and hiring a composer or sound designer was never considered. Also, memory for early game consoles was extremely limited, and the focus was entirely on making graphics. Sound was a mere afterthought. Although the development of audio design was sluggish compared to the visuals (this was because sound technology was primitive and expensive, and seen as the lowest priority), there was certainly an evolution in audio technology. In 1978 the game *Space Invaders* set a new standard with more varied ‘alien’ sound effects, and an audio track featuring a thumping tempo. A year later, Mattel introduced their *Intellivision* system to the market, which contained a sound generator capable of three-part harmony.<sup>15</sup> *Intellivision*’s sports game *Major League Baseball* was also the first game to feature a computer-generated voice. The vocabulary was extremely limited, only consisting of basic words such as “strike”, “ball”, “safe” and “out”, but still served as the first instance of speech in video games, and set the stage for spoken commentary or “play-by-play” in current sports simulator games.

In 1980, *Pac-Man* established some memorable musical elements with its opening theme, the first, and still one of the only, video game melodies that resonated with pop-culture audiences, and was the instigator of melody being used to complement a mood in a video game. *Pac-Man* was the video game industry’s first huge marketing success, with cartoons, clothing, toys, lunch boxes, and other merchandise with the *Pac-Man* logo prominently featured. *Pac-Man* is still one of the most popular video games of all time, and its audio jingles have left an imprint on our culture. The opening theme is timeless, the ‘blinking out’ sound of *Pac-Man* dying is a

---

<sup>14</sup> Winter, David. 1996. "Pong-Story." <http://www.pong-story.com/rhbaer.htm>.

<sup>15</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.3

universally accepted “defeat” signal, and its trademark sounds of Pac-Man eating of dots has been pinpointed as an auditory symbol of “consumerism run amok.”<sup>16</sup> “Reward music” began to be incorporated into levels, sounding a triumphant tune when a level was completed, or a discordant tune when a player failed to finish an objective. The early 80s also saw the release of Atari’s 5200 platform which featured a dedicated audio processor called the Pokey chip. The Pokey chip was a significant breakthrough in terms of game music texture or ‘timbre’. It generated sound using four separate channels, for which the volume, pitch, and distortion values of each could be controlled individually, creating the first four-piece virtual band.<sup>17</sup> Atari’s *Tempest* arcade game had a sound schematic that used two Pokey chips to their full potential, combining for a total of eight separate “voices” that could be arranged in multiple combinations. Atari released a separate soundtrack for *Tempest*, which is identified by the online video game museum *I.C. When* as the video game industry’s first stand-alone audio soundtrack.<sup>18</sup> The 1980s also saw the first integration of human voices and stereo sound in video games in the Cinematronics release *Dragon’s Lair* in 1983. Also in 1983, *Spy Hunter*, which also featured stereo sound, had a channel dedicated specifically to game sounds such as machine guns, helicopters, and other in-game action sounds, as well as borrowing from the well-known *Peter Gunn* theme.<sup>19</sup>

1985 marked the dawn of the Nintendo Entertainment System (NES) console. The home gaming industry boomed with the release of Nintendo’s *Tetris* and *Super Mario Bros.*, which fed the collapse of Atari, and revolutionized the video game industry forever. The 8-bit system used a Motorola 6502 processor, and five monophonic channels. By 1989 one in four people owned one, and by 1996, Nintendo had 6% of the world market share in video games.<sup>20</sup> *Super Mario Bros.* (SMB) set the new standard in game sound, with a full electronic, musical score and more sound effects than any home system had ever used before.<sup>21</sup> Composer for *Super Mario Bros.*, Shigeru Miyamoto, orchestrated the music and sound design by shifting tones to match the action onscreen, therein achieving a new kind of synthesis with gameplay. Sound became inseparable from the game itself, as the sound cues became an integral part of the playing style, which moved video game sound design in new directions, away from cinematic conventions. For example, the pitch would become higher in frequency, and the game music tempo would accelerate to alert the

---

<sup>16</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001. [http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>17</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.3

<sup>18</sup> Thomas Jr., Donald A. "1980s" *PC & Game History*. August 8, 2002 <http://www.icwhen.com>.

<sup>19</sup> McDonald, Glenn. *Gamespot: Brief Timeline of Game Music*. 2001. [http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>20</sup> Research and Development of the US Dept of Energy, 1981. <http://www.osti.gov/accomplishments/videogame.html#pagetop>.



player of time constraints, proximity to an enemy, or of “immunity power-up wearing off.”<sup>22</sup> Sega and Atari reacted by releasing the *Master System* and the *7800 System*, respectively. Both were also 8-bit consoles with similar sound technology to the NES, but paled in comparison to Nintendo’s success, because they lacked a market-breaking hit like *SMB*.

In Japan, 1987, new advances were being made which would impact the North American market a few years later. The first peripheral disk drive for the Nintendo Famicom (the Japanese NES) signaled the beginning of the shift from cartridges to CDs. The Square company also released *Final Fantasy* in Japan, which would become one of the most popular video game franchises to date, generating what is still considered to be some of the best video game music ever made. Composer for *Final Fantasy*, Nobuo Uematsu, broke new ground by integrating cinematic musical scores which had never before been incorporated into video games.

Back in the U.S. in 1989, Nintendo introduced their handheld console, the *Gameboy*, which featured four channels of stereo sound (panning from the left and right speakers). Sega countered with the *Genesis* system, a 16-bit system with six-channel stereo sound. NEC and SNK added some competition with their *TurboGrafx-16* and *SNK Neo-Geo*, respectively. The Neo-Geo featured an 8-bit sound processor with 15 separate channels, while the TurboGrafx-16 had a portable CD attachment that played CD-stored games. This was, in turn, rivaled by Nintendo’s *Super NES* in 1991, which used an 8-bit Sony SPC700 sound chip with 8 separate channels of stereo sound.<sup>23</sup> Sega established itself as Nintendo’s primary competitor, releasing their superior CD-based system entitled *Sega CD* in 1992, and subsequently, two years later, the *Sega 32X*, which added two more sound channels. Nintendo rivaled Sega with the release of the *Nintendo 64* in 1995-1996, a 64-bit machine that was capable of more than 10 times the processing capability of the average Pentium personal computer (PC) at the time.<sup>24</sup>

During the home console war between Nintendo and Sega, the PC began to show promise. Sound quality was beginning to evolve far beyond the limitations of the tinny-sounding internal speaker. Sound cards were being developed and implemented with built-in synthesizer chips allowing for very small message files containing “triggers”, which told the devices what sounds to play and when. PC sound banks consisted of 128 sounds, and were capable of playing

---

<sup>21</sup> Pellis, Jay. "The console Gaming Phenomenon." 1997.

<http://www.geocities.com/sbobbitt21/audyssey24/console.html>

<sup>22</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001.

[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>23</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001.

[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>24</sup> Kline, Stephen. "Killing Time? A Canadian Meditation on Video Game Culture," in *Children and Media Violence*. UNESCO Clearinghouse Annual, 2000.

16 notes at once.<sup>25</sup> The computer sound file began to use an algorithm to compress recorded sounds into ‘.voc’ format for playback, which is similar to today’s ‘.wav’ and ‘.aiff’ files.<sup>26</sup>

Although Nintendo and Sega continued to one-up each other by adding on extra sound channels, synthesizer chips, memory, and digitized voices, video game audio was truly revolutionized with Sony’s introduction of the *Playstation* in 1995. Although the *Playstation* would remain inferior to the *Nintendo 64* with regard to its powerful 64-bit central processing unit (CPU), the 32-bit *Playstation* provided CD-quality stereo sound with “built-in digital support for digital effects such as reverb and looping.”<sup>27</sup> Sony’s technology allowed for more realistic sounding samples and stereo playback, and with the capabilities to store 24 sampled voices, it was feasible for composers and designers to attempt musical scoring. Even more impressive, Sony’s flexible design allowed for three types of music in console games: MIDI, MOD (Digital Module), and Redbook audio.<sup>28</sup> MIDI (Musical Instrument Digital Interface) allowed composers to produce music using a universal format for synthesizer scores. MODs allow for high quality, digital audio samples to be recalled and played back with ease, using short recorded samples. Redbook audio is simply the Red Book standard conformed to for maximum CD quality sound. *Redbook Audio* began to gradually take the place of General MIDI for recording and streaming audio into games, therein giving game sound more of a vivid movie soundtrack feel, using Pulse Code Modulation (PCM) to compress stereo tracks using CDs.<sup>29</sup>

Sega took until 1999 to rival Sony with their highly anticipated, 128-bit, 64-channel, CD-based console, the *Dreamcast*. This was only to be outdone by Sony’s release of the *Playstation 2* in October of 2000, which boasts 48 channels of sound and 2MB of memory dedicated to game audio, in addition to its 128-bit “Emotion Engine” CPU.<sup>30</sup> The *Playstation 2* also has another edge, in that it offers the ability to play DVDs, CDs, and link to the Internet. The PS2 features a main processing unit called the Emotion Engine (EE), powered by a CPU with a capacity for 128 bits and supported by 32MB of RAM. Using DVD storage and streaming allows for higher quality audio, longer loops, and a vast sound bank of samples and sound effects that are stored directly in RAM. Sound processing for the PS2 exists independently, using an IOP (input/output processor), which allows for 2MB of memory to be devoted exclusively to sound, permitting the

---

<sup>25</sup> Marks, Aaron. [The Complete Guide to Game Audio](#). CMP Books, Kansas: 2001. p.3

<sup>26</sup> Ibid.

<sup>27</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." [http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>28</sup> Pidkameny, Eric. "Levels of Sound" *Video Game Music*. 06/15/2002. [www.vgmusic.com/information/vgpaper2.html](http://www.vgmusic.com/information/vgpaper2.html).

<sup>29</sup> Summers, Jodi. *Interactive Music Handbook...* New York: Allworth Press, c1998. p.249

<sup>30</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001. [http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

highest quality sound samples of up to a 48kHz sampling rate, also making it possible to stream continuous background music and ambient soundscapes simultaneously, both in stereo or surround sound.<sup>31</sup> However, Nintendo rebounded, embracing the disk format with the *GameCube* in 2001. Nintendo's *GameCube* supports CD-quality streaming audio, as well as "a programmable digital signal processor (DSP) that supports more than 100 voices and up to 64 simultaneous real-time 3D voices."<sup>32</sup> The beauty of the *GameCube*, from an audio standpoint, is that Nintendo has incorporated a music tool called *MusyX*, which can handle all of the needs for creating audio for a game title, eliminating the need for a development kit, and facilitating sound effects and interactive music creation and sequencing.<sup>33</sup> The latest and most intimidating newcomer to the home console scene is Microsoft, with the release of their *Xbox* console in 2001. With their enormous budget, Microsoft lives up to its name as a computer giant, with a Pentium III class, 733mHz CPU, 64MB RAM, DVD, and an 8GB external hard drive, as well as an Ethernet port to link with both DSL and Cable Internet connections.<sup>34</sup> In terms of audio, Microsoft also has an edge in that it will ultimately be capable of supporting Dolby Digital surround sound, as well as having the capability to broadcast 256 simultaneous stereo voices through 64 different channels, which will be far and above the sound capabilities of both the *Playstation 2* and *GameCube*.<sup>35</sup> The *Xbox*'s unit for audio data is the Media Communications Processor (MCP), which includes four independent audio processors for digital audio and 3D positioning: the set-up engine, Voice Processor (VP), Global Processor (GP), and Encode Processor (EP).<sup>36</sup> The processors each perform a specific function to ensure the highest audio quality, and in conjunction with DirectMusic, DirectX 8 and Windows Media software tools for creating and implementing game sound, audio scripting and programming become relatively simple, and are compatible cross-platform from the *Xbox* console to the PC. As we will see in greater detail in chapter three of this thesis, the audio hardware and software built into the *PS2*, *GameCube*, and *Xbox* make interactive and adaptive audio truly possible.

---

<sup>31</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.275

<sup>32</sup> Satterfield, Shane. "Gamespot Presents: GameCube Dossier". *Gamespot*. May 10, 2002. [http://gamespot.com/gamespot/features/video/gamecube\\_dossier/p2.html](http://gamespot.com/gamespot/features/video/gamecube_dossier/p2.html).

<sup>33</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.276

<sup>34</sup> Pidkameny, Eric. "Levels of Sound" *Video Game Music*. 2002. [www.vgmusic.com/information/vgpaper2.html](http://www.vgmusic.com/information/vgpaper2.html).

<sup>35</sup> Gamespot.com: "Gamespot Presents: Xbox Dossier." 2002.

[http://gamespot.com/gamespot/features/video/x\\_box\\_dossier/p1\\_01.html](http://gamespot.com/gamespot/features/video/x_box_dossier/p1_01.html).

<sup>36</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.279

### 1.1.3 At Present: "Hear We Are"

Now, with Microsoft getting in on the action with the *Xbox* console, Nintendo with its new and improved *Gamecube*, Sony's revamped *Playstation 2*, and a wide array of high quality PC and online computer games, there is a new standard in home video game technology. Rich, sophisticated, orchestral scores by composers are now becoming an industry standard, and a common expectation from avid gamers. Channel and voice capabilities, alone, have progressed so rapidly in the past few years that one can only imagine what's to come. We are in the midst of another gaming boom, and this time the technology is advancing in leaps and bounds like never before. Graphics are becoming so real, that sometimes it is difficult to distinguish digital characters from photographic images on the screen. What was thought unimaginable half a decade ago is now becoming a reality with realistic, vibrant game graphics and vivid audio.

Instead of soundtracks using *MIDI* and onboard wavetable synth formats (that use built-in synthesizers and soundcards), game music is now composed using *Redbook* audio and digital modules (MODs) that stream customized audio samples.<sup>37</sup> Much attention is now being devoted to interactive audio considerations, where audio reacts to gamestates and player actions (as I will discuss in chapter three), with the rise of DLS (Downloadable Sample) technology, and adaptive audio engines. The quality has improved to a great extent, with all consoles now using CDs, all capable of transmitting digital stereo sound, 44.1 kHz sampling rate, and 16-bit audio tracks; in essence, full CD-quality sound. The CD, along with the console's CPU, also has the capacity to hold an abundance of audio and visual information, in addition to memory, which allows for more detailed speech, music, sound, and graphics. Many of the newer games now have an online component, expanding the game far beyond home use on individual consoles. Trends in online gaming are spreading like wildfire, with huge, global gaming communities being created on the Internet, countless websites devoted to game audio and music, and worldwide discussion forums and websites devoted to video game dialogue, feedback, and strategies.

Game audio has advanced to incorporate interactive, gestural, epic musical scores and complex soundscapes that are now compared to award-winning motion picture scores. In 2000, video game music reached a previously unthinkable milestone, by gaining recognition from the National Academy of Recording Arts and Sciences (NARAS). Video games are now eligible to compete for annual Grammy awards against its predecessors, TV and film. Game soundtracks can now be considered in three general categories: Best Soundtrack Album; Best Song; or Best

---

<sup>37</sup> Brandon, Alexander. "Interactive Music: Merging Quality with Effectiveness." 1998.  
[http://www.gamasutra.com/features/sound\\_and\\_music/19980327/interactive\\_music\\_quality\\_intro.htm](http://www.gamasutra.com/features/sound_and_music/19980327/interactive_music_quality_intro.htm)

Instrumental Composition for a Motion Picture, Television, or Other Visual Media.<sup>38</sup> It is a huge leap for audio in this interactive medium to receive such high recognition, when it is still so often dismissed by most as low-brow entertainment. Game audio has now become an art form of its own. With heightened audio quality, game music CD soundtracks competing for music store shelf space, and Grammy award consideration and potential, game audio composers, programmers and designers are on an equal level with those of the television and film industries.

The amount of choice and variety in video games is now plentiful. There is much competition for the consumer's dollar and pressure to conform to video game players' preconceived notions of how a game should function, sound, and look, while adding some new, attractive features to add novelty. Today's gamers demand more than industry standard. Games range from all different genres and types. There are games aimed at children, as well as adult audiences. Games borrow from TV, film, and other media, or they set the trend for other media to follow. The video game medium, because it is still relatively new, flexible, inclusive of other forms of media, and, first and foremost, *interactive*, it can push the envelope, and tread on new technological ground. That is why music, TV, and film industries, as well as advertising giants, are flocking to this new medium to reach their target audience. Let us now examine how interactivity factors into the video game equation.

## 1.2 Interactivity & Immersion

This section aims to introduce concepts of interactivity, from its beginnings in cybernetics, to its current role in interactive home video gameplay. Up to now we have looked at the background of video games as interactive media, but the term 'interactive' has been used loosely, and has yet to be adequately defined. Starting with a general definition of what constitutes interactivity before applying it to video games might be the best way to achieve an understanding of how interactivity applies to the interactive and adaptive audio concepts being introduced in this thesis.

In the simplest and most general form, to interact means to “act one upon another”,<sup>39</sup> to engage with another person or object. Human interactivity, of course, is the basis for any community or society in general. Theories of interaction with machines, the nature of the interactivity being referred to in this study, date back to the earliest discussions of multimedia in the 1950s: Cybernetics. In 1954, Norbert Wiener of MIT coined 'cybernetics' as the science of

---

<sup>38</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001.  
[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>39</sup> Webster. *Webster's Encyclopedic Unabridged Dictionary of the English Language*. Dilithium Press, Portland House, NY: 1989. <Interact> p.739

messages transmitted between man and machine. Cybernetic events range from action-response cues such as driving a car, clicking a mouse, or passing through an automatic door.<sup>40</sup> Wiener's premise was that human communication should be the model for all human-machine exchanges, and that interactivity between humans and computers is affected by multiple factors, namely entropy, control and, most importantly, the principle of feedback.<sup>41</sup> He suggested that interactivity entails an order being sent out, followed by a signal of compliance in exchange, and that this dialogue becomes a symbiotic partnership, regardless of if the exchange takes place through a machine rather than direct human communication. In his book *The Human Use of Human Beings*, originally published in 1950, Wiener defines the fundamental principles of the key element of feedback within Cybernetics, classifying communication and control together as interchangeable rules of interactivity that are integral to all forms of human communication and interaction. Wiener states that "feedback tends to accomplish the purpose of... voluntary feedback, for in man we consider that a voluntary action is essentially a choice between tropisms."<sup>42</sup> This is to say that communicating a message to another person contains information which is accessed and therein gives control to the receiver as to how that information will be processed, obeyed, and understood. Whether this transmission of information is conducted through a machine is irrelevant.<sup>43</sup> It was Wiener's insight that served as the premise for all human-computer activity and interface design to follow, and was the basis for all information-handling machines we know and use today. He saw the potential for an interactive dialogue between humans and machines, therein creating a cooperative, working relationship between them. Inevitably, debates emerged. There were optimists who romanticized the interactive, cybernetic relationship, claiming computers could solve the problems of the world (Engelbart, 1962). Nicholas Negroponte, author of *Being Digital*, writes from an optimistic angle about the potential for computers and humans to form a "new, global social fabric", and that non-linear technological shifts toward interactive user-machine relationships could serve to empower and accelerate human capabilities.<sup>44</sup>

It was Myron Krueger who explored computers as having the capacity to combine arts and sciences into an interactive art form. In 1970, Krueger created real-time computer mediated

---

<sup>40</sup> Licklider, JCR; Packer, Randall; Jordan, Ken, (eds.) *Multimedia: From Wagner to Virtual Reality*. New York : Norton, 2001. p.48

<sup>41</sup> Ibid. p.48

<sup>42</sup>Publius, "Norbert Wiener Quotes" (taken from *The Human Use of Human Beings: Cybernetics and Society*. 2<sup>nd</sup> ed., 1954) <http://www.angelfire.com/co/1x137/wiener.html>

<sup>43</sup> Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society*. 2<sup>nd</sup> ed. Garden City, NY: Doubleday, 1954.

<sup>44</sup> Negroponte, Nicholas. *Being Digital*. London: Hodder & Stoughton, 1995. p.11-13

spaces, entitled “responsive environments,”<sup>45</sup> in which he explored ‘telematic art’ to compose video-generated computer environments that were programmed to respond and even anticipate audience reactions. Krueger, like Negroponte, saw us as moving toward complete immersion in a computer-created world. He saw humans and machines as becoming intertwined, and more than a vehicle or tool, he saw the potential for a partnership between them. Krueger’s idea in his project *Videoplace* was to create a responsive environment that would be a thinking machine with the ability to respond intelligently to the actions of the audience. His goals were almost prophetic, in the sense that many of his plans fit the profile for the current state of interactivity in video games.

Over a period of time the computer displays establish a context within which the interaction occurs. It is within this context that the participant chooses his next action and anticipates the environment’s response. If the response is unexpected, the environment has changed the context and the participant must reexamine his expectations. The experience is controlled by a composition which anticipates the participant’s actions and flirts with his expectations.<sup>46</sup>

The concept was to create an artificial context, in which the artist maintains complete control over the laws of cause and effect, and the participant would have control over the course of actions within the alternate reality. The model sounds vaguely familiar to Sony’s slogan for the *Playstation 2* home console, which also urges the gamer to escape into an artificial reality world: “Live in your world. Play in ours™.”<sup>47</sup> Media artist Lynn Hershman, in her article “The Fantasy Beyond Control,” suggests that interactivity, as the marriage of computers, text, image, and sound, challenges the notion that viewing art is necessarily a passive activity. She claims that interactivity requires viewers to react and make choices, and that interaction with virtual characters in a virtual space actually encourages audience members to break free from a conventional audience role, to navigate through narrative paths instead of sitting back as a passive recipient (Hershman, 1989). Implications for an interactive art form are that the participant provides the real-time authorship for the experience.

Historically, the notion of interactivity as applied to media audience perception has also given rise to a number of debates surrounding the individual’s ability to interact with media messages and content. For Adorno, Horkheimer and Marcuse (three primary Frankfurt School

---

<sup>45</sup> Krueger, Myron; Packer, Randall; Jordan, Ken, eds. *Multimedia: From Wagner to Virtual Reality*. New York : Norton, 2001. p.105

<sup>46</sup> Ibid. p.106

<sup>47</sup> Playstation.com. <http://www.us.playstation.com/games/SCUS-97134.asp>. © 2003 Sony Computer Entertainment America Inc.

theorists), to analyze a work of art or 'cultural artifact' is to "analyze and assess the way it is interpreted. This entails an inquiry into its formation and reception."<sup>48</sup>

One of the most significant successes of the Frankfurt school was the shaping of electronic media into the "powerful instruments of social control which they represent today."<sup>49</sup> The Frankfurt School mentality is in large part building and continuing on the Marxist concepts of *commodity* and *fetishism*. Adorno and Horkheimer claim that the structure of mass society doesn't allow for adequate response, and actually erodes the capacity for people to make independent judgements (Adorno & Horkheimer, 1944). This translates as a source of unreason for Frankfurt School theorists, in that people think as a collective, mass audience, and are merely given the 'illusion' of choice, essentially liquidating the individual.<sup>50</sup> Ted Friedman, in his article *Making Sense of Software: Computer Games and Interactive Textuality*, extends Adorno's insight, relating the discussions specifically to interactive video games. Friedman presents the interactive nature of video games as a paradox, asking "How can one give the player a sense of 'control' over the game while still propelling the player through a compelling narrative?"<sup>51</sup> Friedman continues along the same lines as Adorno, claiming that interactive games cannot allow for autonomy, but merely offer the player a series of choices leading to other choices, which ultimately leads to a limited set of predetermined options laid out by the game designing team, therein giving players an *illusion* of agency (Friedman, 1975). Horkheimer and Marcuse identified the process of media reception as the "collapse of the autonomous individual", commenting on mass media as transmitting the 'required' values which leads to the 'manipulated consciousness' of man, therein contributing to man's ignorance of the world in which he lives.<sup>52</sup> Adorno and Horkheimer continue on this topic of commodities with a detailed discussion.

Consumers appear as statistics on research organization charts...the technique is that used for any type of propaganda...there is nothing left for the consumer to classify. Producers have done it for him...the object is to overpower the consumer, who is conceived as absent-minded or resistant.<sup>53</sup>

---

<sup>48</sup> Held, David. *Introduction to Critical Theory*. London: Hutchinson, 1980. p.77

<sup>49</sup> Minnicino, Micheal J. *The New Dark Age: The Frankfurt School and "Political Correctness."* Courtesy of the *Schiller Institute*. From the Winter, 1992 issue of *FIDELIO Magazine*. [http://www.schillerinstitute.org/fid\\_91-96/921\\_frankfurt.html](http://www.schillerinstitute.org/fid_91-96/921_frankfurt.html)

<sup>50</sup> Arato, Andrew; Gebhardt, Eike. "On the Fetish-Character in Music and the Regression of Listening" *The Essential Frankfurt School reader* (Chapter by Theodor W. Adorno). NY: Urizen Books, c1978. p.277-280

<sup>51</sup> Berger, Arthur Asa. *Video Games: A Popular Culture Phenomenon*. Transaction Publishers, NB, London: 2001. p.81

<sup>52</sup> Swingewood, Alan. *The Myth of Mass Culture*. London : Macmillan, 1977. p.14-15

<sup>53</sup> Adorno, Theodor, Horkheimer, Max. *Dialectic of Enlightenment*. Translated By John Cumming. New York: Herder and Herder, 1972. p.123-125,163



Adorno further solidifies his views in the following statement.

The masochistic mass culture is the necessary manifestation of the almighty production itself... It corresponds to the behavior of the prisoner who loves his cell because he has been left nothing else to love.<sup>54</sup>

This is a very strong statement, emphasizing a very pessimistic view of society. In essence, Adorno is stating that we, as a commodified mass society, are actually *imprisoned* by our relationship to the system that encompasses (and contains) us.

Because of the supposed “passivity” of the masses, Adorno gives no credit to individuals to interpret and decode what’s being perceived. In fact, he says that the nature of leisure and pleasure constitutes not having to think, and not having to resist. While it may be true that entertainment is often seen as an escape that doesn’t require much thought, video games surely are exempted from this rule, as audiences expect and want to be stimulated in their entertainment practices. Video games, although often dismissed as low brow entertainment, can actually be viewed as an exception to the passive media model, as video game playing demands full immersion and engagement from an active participant, and cannot possibly be seen as a worthwhile activity unless it provides multiple levels of interactive stimulation.

In The Media and Modernity, John Thompson speaks of the development of modern societies in terms of how technology changes forms of interaction and communication among individuals. He examines the way in which interaction is made possible by new communication media and how the development of media technology gradually alters the nature of tradition and individual reactions to it. Thompson urges us to look at the way in which we use media products, and suggests that we examine possible reactions to the messages we receive (Thompson, 1995). New levels of interactivity made possible by technological advances in video games certainly serve as a means to re-examine the media reception and interaction model. Author George E. Lewis, writing for *Western Front* magazine, discusses interactivity in media and the need to discover and redefine new levels of engagement in the following excerpt.

... If a imaginatively stimulating and communally informed experience, such as live music performance, is to be considered “passive”, while a consumer sitting alone in front of a screen, provided with a set of choices to navigate between, is considered “interactive”, then perhaps the real passivity consists in the impotence that we must all feel in not challenging the power that has redefined, according to its own needs, what interactivity means and when and how we will engage in it.<sup>55</sup>

---

<sup>54</sup> Arato, Andrew, Gebhardt, Eike. "On the Fetish-Character in Music and the Regression of Listening" The Essential Frankfurt School reader (Chapter by Theodor W. Adorno). NY: Urizen Books, c1978. p.277-280.

<sup>55</sup> Lewis, George E. "Singing the Alternative Interactivity Blues." *Western Front magazine*, and the Association for the Advancement of Creative Musicians website: 1995 <http://audiolabo.free.fr/revue1999/content/lewis2.htm>.

Lewis expresses some negativity about 'choices' offered by interactive media, claiming that interactivity as offered by media producers is simply a means to exert control over the consumer through a means to predetermined outcomes, therein providing the illusion of an autonomous, selective process.

J.C. Herz takes the opposite stance in discussing interactivity in video games in her book Joystick Nation: How Video Games Ate Our Quarters, Won Our Hearts, and Rewired Our Minds. She pinpoints technological developments in video games as taking entertainment to new, interactive levels, becoming like a vivid motion picture in which players have the opportunity to participate.

What makes [a video game] immersive is a world where no territory is off-limits, anything you see is fair game, and all your actions have consequences. This is what game designers call a "real-time object-oriented environment," which is to say time moves ahead and the world churns even in places you don't care to look (or haven't found). Characters exist independently. Options shift. Events -some completely beyond your control- unfold in a world that can age... with each tick of its internal clock, the fictional world changes. (1997, p. 155)<sup>56</sup>

Furthermore, Janet H. Murray, discussing the future of narratives through digital media in Hamlet on the Holodeck, describes how hypermedia has finally opened our eyes to the active role audience members take in engaging with a story.

Just as we have only recently learned to think of the solitary reader as playing an active role through imaginative engagement with the story, so too are we just beginning to understand that the interactor in digital environments can be the recipient of an externally authored world.<sup>57</sup>

Positive or negative, the notion of interactivity is based on the concept of non-linear thought. Plain and simple, human thought is not a linear process, and that is why media are gravitating more and more toward becoming interactive, because an entertainment experience in which the participant feels he or she has influence over the outcome becomes a more engaging activity. Media interactivity can be described as "the ability of the user to interact with a computer which can be used to control the flow, pace, and content of a program."<sup>58</sup>

It is important to note that not all interactivity (in media or in the natural, acoustic realm) has the same value, and it would be naive to assume that any process that involves some element of choice constitutes a comparable interactive exchange. While it could be called interactive, in

---

<sup>56</sup> Berger, Arthur Asa. Video Games: A Popular Culture Phenomenon. Transaction Publishers, NB, London: 2001. p.9

<sup>57</sup> Murray, Janet H. Hamlet on the Holodeck... The MIT Press, Cambridge, 1997. p.275

<sup>58</sup> Summers, Jodi. Interactive Music Handbook: The Definitive Guide to Internet Music Strategies, Enhanced CD Production, and Business Development. New York: Allworth Press, c1998. p.247

the simplest form, to say that the act of turning on a television set or pushing buttons on a remote control constitutes a media interaction, it must be understood that interactive audio for video game environments, the subject of this thesis, involves far more of an interactive partnership made possible by complex, algorithmic audio processes.

Interacting with soundscapes is not a new discovery, as it can be traced back to the earliest and most basic ear-cleaning exercises in soundscape studies. When a listener goes on a listening walk or *soundwalk*, which is an exploration of the soundscape for the sake of ear cleaning (Schafer, 1977), the freest degree of interactivity transpires between the listener and the natural, acoustic environment. While machine mediation constrains the full potential for an interactive relationship between the listener and the soundscape, complex algorithmic systems, like those inherent in interactive audio for video game console systems, present user with a number of pathways and outcomes to choose from. The idea here is that a linear decision tree does not constitute an interactive exchange for the purposes of deconstructing audio interactivity in this thesis. The vast number of possible outcomes and variables, made possible by real-time, interactive audio technologies in modern video games, creates an interactive partnership through which a feedback cycle is created between human and machine, as put forth in the theories of Cybernetics (Wiener, 1954). As will be seen later in the thesis, the vivid sonic environments created by designers as virtual playgrounds for the player to engage with reactive auditory elements begin to resemble the interactive engagement intrinsic to aural soundscape traditions. The high potential for interactivity within the video game construct becomes heightened when examining new methods of designing and implementing interactive audio components.

The concept of toys linked with television for a new interactive form of play was introduced in 1987 with the *Captain Power* game, which allowed children to participate in the act of computerized play that they were previously only able to watch. This allowed viewers to engage with an actual plot, narrative or storyline. Games like *Captain Power* gave children the capability to practice fighting roles as an extension of the TV show, which they could interact with and play long after the show came to an end. This new level of interactivity can be seen in modern quest games such as the *Final Fantasy* or *The Matrix* universe, which have subcultures of their own, enveloping gamers into virtual, fantasy realms. The newest movement in modern home video games is a peripheral device (an accessory made available by the three major console system giants: Sony, Microsoft, and Nintendo) that can be purchased separately, which links the home console to the Internet, allowing the home gamer to associate with an unlimited global universe of virtual gaming on the World Wide Web. Now, not only do interactive technologies play out within the home consoles themselves, but players have the opportunity to battle with or

against other unpredictable, spontaneous human counterparts across the globe, all united into a gaming community on the *Information Superhighway*.

The audio component in home video games is a driving force in making games into the interactive and potentially immersing experiences they have become. In 1978 Midway introduced *Space Invaders*, which had no music per se, but featured a thumping, rhythmic tempo and sound effects that would accelerate as the enemies got closer. This is seen as the first example of non-diegetic sound in video games, which is referred to in film as "background music amplifying the mood of a scene and/or explicating the dramatic developments and aspects of character."<sup>59</sup>

The goal for game designers is to make video games as engaging and immersive as possible to keep up with the demands of the consumer, which requires making the virtual fantasy realm of each game into an interactive, participatory environment through which the player has a sense of agency over his or her own fate and destiny. The notion of *immersion* in this thesis refers to the experience of being transported into richly-constructed fantasy environments in which players can act out their fantasies, which provides a sense of satisfaction for the user in being submerged into an alternate reality. Immersion in digital media, as put forth by Janet H. Murray, is used as a metaphorical term "derived from the physical experience of being submerged in water...the sensation of being surrounded completely by another reality...that takes over all of our attention, our whole perceptual apparatus."<sup>60</sup> The idea is that video game environments serve as fictional worlds that are traversable and interactive, intensified by the participatory nature of the gaming experience. Furthermore, Murray maintains that because of the players' desire to experience immersion, users focus attention on the all-encompassing world created by game designers, and use their intelligence to validate rather than to question the realism and authenticity of the experience. She uses the metaphor of an "immersive visit" to describe the interactive participation inherent in surrogate environments, in that experiencing multisensory immersion is structured by establishing borders between the virtual world and ordinary life. The idea is that video games, while providing a fictional environment with goals to accomplish and telling an immersive story, provide the structure for an immersive visit through a "densely textured visual and aural environment... [providing for] intense immersion."<sup>61</sup>

Murray suggests that the pleasurable immersion inherent in being submerged within electronic environments is partially brought about by a sense of *agency* for the player. Agency within a narrative environment, Murray explains, is "the satisfying power to take meaningful

---

<sup>59</sup> Macy, Laura (ed.) "Film Music." *Grove Music*. 2<sup>nd</sup> ed. The New Grove Dictionary of Music and Musicians. May 12, 2002 <http://www.grovemusic.com/index.html>.

<sup>60</sup> Murray, Janet H. *Hamlet on the Holodeck...* The MIT Press, Cambridge, 1997. p.98

<sup>61</sup> *Ibid.*, p.108-109

action and see the results of our decisions and choices."<sup>62</sup> The idea here is that the act of navigating through the video-game narrative serves as a meaningful experience that provides an element of control or agency for the player over his or her own destiny within the gaming context. The combination of unexpected surprises and chosen outcomes serves as "an open road with wide latitude to explore more than one way to get somewhere... using the act of navigation to unfold a story that flows from our own meaningful choices."<sup>63</sup> This is the premise behind interactivity in the digital medium of video games, that pleasure is derived as the player "encounter[s] a world that is dynamically altered by [the user's] participation."<sup>64</sup>

The fundamental differences between audio in fixed media, such as film and television, and that of interactive entertainment media, namely video games, is that music and sound in interactive entertainment can react to the users, thereby drawing participants into the game world that has been created for them. Audio has the capability to identify with a particular audience, and can instantly append associations, symbols, and definitions beyond what the visuals can achieve. The interaction is the basis for making a video game immersive, entertaining, and successful, and it is therefore only natural for designers to want to blend the interactive control found in video games with the immersive qualities of music and ambient sound.<sup>65</sup>

As we will see in greater detail in the third chapter of this thesis, there are different levels and technological applications of interactivity in video game audio. Practitioners make the distinction between reactive and adaptive audio as varying forms of audio interactivity. Some audio scores are made up of predetermined cues or scripts that change according to environmental aspects in the game, referred to by some as *adaptive* audio rather than interactive, because it is audio that responds to transitions in the game rather than reacting specifically to a player's given actions. Before we attempt to deconstruct interactive audio in video games, we must first understand the fundamentals of the acoustic parameters and audio environments or *soundscapes* that serve as the auditory foundation for the bulk of this study. By understanding some of the acoustic parameters at play, how they function in soundscape studies, and how sound components interrelate and juxtapose within a given context, we can then move on to analyze how they can apply to an interactive, electroacoustic virtual environment, which is the framework for any video game soundtrack. Only through the exploration and deconstruction of sound, on an informational level, can we begin to comprehend the pertinence of interactive audio considerations and the

---

<sup>62</sup> Ibid., p. 126

<sup>63</sup> Ibid., p. 132-133

<sup>64</sup> Ibid., p. 129

<sup>65</sup> Patterson, Scott. *Interactive Music Sequencer Design*. Gamasutra, 2001: [www.gamasutra.com/resource\\_guide/20010515/patterson\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/patterson_01.htm).

potential for it to have a direct impact on the user. Once this is examined, I can then proceed to further compose the analytical categories to be applied later in this study.

*Can music, sound effects, and narratives exist simultaneously in the soundscape without negatively impacting the game experience? Yes, they can, but a certain amount of care needs to be taken to ensure they do...what ultimately matters when doing any audio for video games is that everything blends... [keeping] a game's audio sounding cohesive.<sup>66</sup>*

-Aaron Marks  
Video Game Audio Designer

## 2. Sound: An Ear-Opening Experience

This chapter aims to introduce basic acoustic terminology integral to a fundamental understanding of the key notions being applied and examined in this study. An overview of the central concepts established in soundscape studies also serves as a paradigm for understanding the elements interacting within the construct of a surrogate auditory landscape. It is also imperative that the reader be introduced to the concepts vital to *psychoacoustics*, an area of study that combines theories of acoustics with psychology. Although psychoacoustic considerations will be mostly applied later in this study, it is also important to gain a basic understanding of the main concepts within psychoacoustics before embarking on analyses of specific games. By reviewing the principal terms at the foundation of acoustic theory, by examining the function of acoustic elements in relation to their environment, and acquiring a basic knowledge of potential psychological impressions evoked by sonic elements, we can then move on to construct the analytical framework to be used in the case studies as textual audio analyses of specific video games and their interactive function. This chapter seeks to inform the reader by illuminating the key acoustic and electroacoustic terms relevant to an understanding of the concepts central to this study, and its scope cannot possibly incorporate more than a basic exploration of acoustic theories, psychoacoustic factors, and soundscape applications vital to this thesis.

### 2.1 A “Sound” Approach

Analyzing video game audio as a media text that creates a surrogate, electroacoustic environment is a new approach with respect to the video game medium. It is a methodology that requires an understanding of acoustics, which entails the physical components of sound, psychoacoustics, and soundscape studies, which is a unique area of study that applies both acoustic and psychoacoustic theories to aural characteristics that make up a sonic environment. The pioneers of soundscape studies, namely R. Murray Schafer and Barry Truax, examined auditory environments through the study of communication. Schafer posited sounds as

---

<sup>66</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.261-262

inseparable from the context in which they are perceived, and felt that, in order to appreciate how sound functions in an 'auditory landscape', one must take into account several contextual sonic factors. According to Schafer, we are only able to comprehend sounds through their function in an environmental context, taking into account the acoustics (what the sounds are), psychoacoustics (how they are perceived), semantics (what they mean), and aesthetics (if they appeal).<sup>67</sup> Truax further developed the study of sound by grouping three areas as *systems of acoustic communication*, determining that sound can be organized and understood by deconstructing *speech*, *music*, and the *sonic environment* (soundscape) on their common foundation as "organized sound."<sup>68</sup> The continuum organizes the three major systems of acoustic communication as a model for human communication as follows:

Speech – Music – Soundscape<sup>69</sup>

The reason for distinguishing between these systems is because the rules and relationships can vary depending on whether acoustic elements are being examined as speech, music, or soundscape. In order to decipher the meaning of sounds and their behaviour within an auditory environment, we must recognize that musical rules, for example, are not necessarily the same as those applied to speech or soundscape. While some acoustic aspects may apply to all three systems of acoustic communication, they may have different implications and may behave differently in relation to which system it is being analyzed within. By analyzing sound in terms of speech and vocal elements, music as organized and composed sound, and the auditory environmental context in which they interact, we can use the continuum as a basis for understanding their sonic relationship and organizational features, and further deconstruct them as systems of acoustic communication. This is the methodology I will be applying for this study.

Using the continuum as a model for functional analysis allows me to analyze video game soundscapes as virtual, surrogate environments. By using the speech-music-soundscape continuum as a model to construct analytical categories, I can then review the specific audio elements at play, and hypothesize about the overall capacity for interactivity. It is important to note that in film practice speech, music and sound effects (SFX) are often segregated, and treated as separate entities. While the deconstruction of sound is a logical approach to understanding the function of each of these systems within an interactive audio construct, it is important to distinguish that, unlike film, the video game medium is one in which the lines between these

---

<sup>67</sup> Schafer, R. Murray. *The Soundscape*, Destiny Books, Vermont: 1994, 1977. p.148-150

<sup>68</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.50-51

<sup>69</sup> Ibid. p.50



acoustic systems become blurred. The sonic elements function very much as a whole to create a surrogate environment, and allow for a truly engaging, interactive experience. First, let us examine some of the general acoustic and *electroacoustic* (the electrification of an acoustic source) terms that can apply to the three systems introduced in the speech-music-soundscape continuum.

Sound occurs when an object vibrates or *oscillates* from a resting position. Probably one of the most basic terms in the study of acoustics is *amplitude* or *intensity*, which refers to the magnitude (perceived as volume or loudness) of a sound over time.<sup>70</sup> How loud a sound is perceived depends on the amount of sound pressure in a given sound. However, the intensity of a sound is generally not static or fixed, and changes over time. The shift in a sound's intensity over the duration of the sound is referred to as the *envelope* of the sound, which is the shape of a sound's intensity before it diminishes to silence. The sound's behaviour over time is measured by singling out distinctive features such as the *attack*, or the initial part of the sound, the *steady state*, or sustained portion of the sound envelope, and the *decay*, which is the release of the sound including any reverberation as it loses amplitude due to loss of energy.<sup>71</sup> Within the envelope, the initial part of a sound is called a sound's *transients*, which is what distinguishes or identifies the sound's source. For example, a good way of understanding onset transients is thinking of a violin's envelope in the musical context. A violin sound will have an attack (when the sound begins), a steady state (as the bow glides along the strings), and a decay (when the sound dies away once the bow is lifted), but what truly identifies that it is a violin sound as opposed to another instrument for the listener is the transient, which is the scratchy initial sound that the bow makes right before the melodic string sound ensues once the instrument is played. In speech, a linguistic example of a transient would be the initial consonant sound that distinguishes a word's significance to the listener in words like 'pill', 'kill', or 'bill'. These words, of course, have completely different meanings that could not be distinguished without the initial consonant sound as the transient to identify the source meaning.

Another component essential to a basic understanding of acoustics is the *frequency* of a sound. Frequency, in the periodic sense, results in the 'pitch' of a given sound, and is determined by the number of cycles of vibration that occur in a second of sound.<sup>72</sup> Sounds are mostly made up of multiple frequencies within a *spectrum*, and the variation of pressure cycles shown on an oscilloscope screen are called *sound waves*. A sound that combines higher frequencies in a given

---

<sup>70</sup> Pellman, Samuel. An Introduction to the Creation of Electroacoustic Music. Wadsworth Publishing, California: 1994. p.5

<sup>71</sup> Ibid. p.5-6

<sup>72</sup> Ibid. p.4

spectrum (which is the frequency content of a sound) is referred to as being 'brighter', while a tone that sounds 'heavier on the bass' seems 'darker', being composed of low frequency sound. There are a wide number of common waveforms within the study of acoustics, the most common ones being the Sine, Triangle, Sawtooth, and Square shaped waveforms.<sup>73</sup> Each of these is so-named for its wave shape as depicted graphically below.

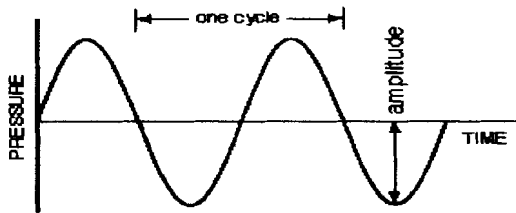


Figure 1: Sine Waveform

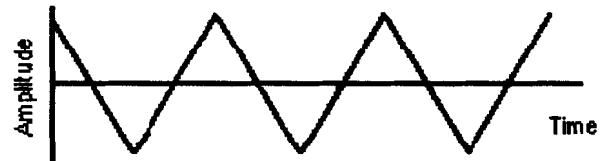


Figure 2: Triangle Waveform

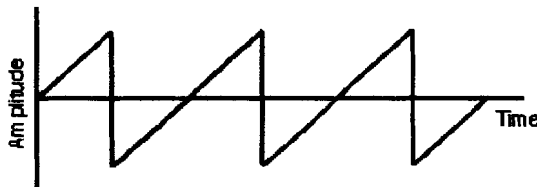


Figure 3: Sawtooth Waveform

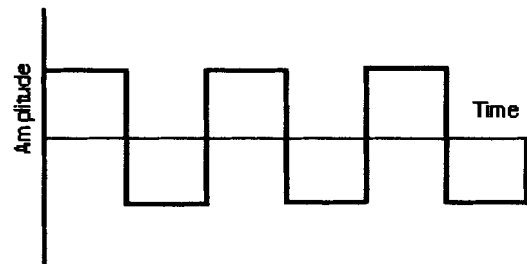


Figure 4: Square Waveform

The more frequencies that are acting at once to make a given sound will change the character of the sound itself, which can impact the tonal quality of a sound. The spectrum of a sound over time (including transients) will alter the *timbre*<sup>74</sup> of a sound, as I described in an earlier example of the violin. This has to do with the perception of combined frequency components that gives sound a specific colour or character that allows us to distinguish between sounds. Timbre is often described as tone colour (e.g., bright, dark) or texture (e.g., harsh, smooth, crisp, dull, sharp, etc.),<sup>75</sup> which is important in identifying a sound's source, as well as giving a sound a unique quality. It may sound simple, but the tone quality or colour of a given sound can be quite complex, as it consists of many simultaneous frequencies mixed together. This fusion is what we perceive as the timbre of the sound.<sup>76</sup> Timbre is a very important element in acoustics, and is of considerable importance with regard to musicians and sound designers using

<sup>73</sup> Ibid. p.233

<sup>74</sup> *Timbre* is a word of French origin referring to quality or tone of instrument or voice, pronounced "TAM-ber".

<sup>75</sup> Truax, Barry. "TIMBRE", Handbook for Acoustic Ecology CD-ROM, Cambridge Street Publishing, 1999.

contemporary electroacoustic technology. Technology has made meticulous control of timbral components possible, which comes into play especially with respect to the richly-fabricated electroacoustic environments evident in modern video games, as we will see later in this thesis.

## **2.2 Speech – Music – Soundscape**

Now that we have gained a basic understanding of the fundamental terms at the root of acoustic terminology, we can move forward to establish more concrete and specific sound elements and phenomena as applied to the three major acoustic communication systems. By focusing on specific aspects of each scheme within the continuum, we can achieve a greater understanding of the rudiments that make up these categories that will serve as the basis for the analytical structure in deconstructing video game audio in the textual analyses to follow. It is important to note that although I am separating these systems for the sake of explaining the individual functions of speech, music, and the soundscape respectively, they are all inextricably linked in the effectiveness of an overall video game soundscape. Game audio designers, as we will see later in this thesis, maintain that all of these elements must be considered together in order to design an effective, interactive sonic environment. As Aaron Marks insightfully pinpoints “the game soundscape is every element that makes sound: the music, the sound effects, and the narratives. It is the purposeful blend of these elements that form an organized presentation known as a “soundscape.” Without this type of order, there is audio chaos.”<sup>77</sup>

### **2.2.1 Speech: More than Just Hearsay**

Speech is vocal sound, which is the primary basis for most human communication. Human speech involves facets of linguistics (the study of words, or verbal communication) and sound-making (paralanguage), coupled with physical gestures such as body and facial movements (known as the study of kinesics), which are commonly regarded as the components that go into vocalization and human expression. I will not delve too deeply into linguistics or kinesics except to make mention of them as pertinent components of speech. Conversely, I do feel it is necessary to mention that linguistics involves the study of verbal communication, or the phonemic and syntactical components of human language, which can be further dissected in a study applying discourse analysis as a research methodology to determine the significance of linguistic patterns in deciphering semantic meaning.

---

<sup>76</sup> Pellman, Samuel. An Introduction to the Creation of Electroacoustic Music. Wadsworth Publishing, California: 1994. p.210

<sup>77</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.292

Rather than focusing on the words themselves, Barry Truax suggests we direct attention to the significance of '*paralanguage*', which is often neglected when studying speech components. Paralanguage refers to the voice quality or 'tone of voice' that contains information through nonverbal, acoustic utterances that convey meaning and personal expression. It is often thought of as the musical aspect of speech, in that how something is said involves such musical qualities as "inflection (pitch contours), rhythm, phrasing, emphasis (or accent), punctuation, timbre (or sound quality), silence (rests), and even cadences", which are all aspects describing melody.<sup>78</sup> Truax feels that paralanguage is perhaps the most pertinent component of speech as human communication, as it "contains essential information for understanding the meaning of the utterance."<sup>79</sup> *How* something is said as opposed to *what* is said allows us to gain insight into the role that human sound-making, and communicates more significant information than the straightforward linguistic content of the words themselves. Paralinguistic aspects depend on dynamics of the auditory system such as pitch, inflection, and stress patterns, which can convey deeper meanings, such as suggesting persuasion by the orator.<sup>80</sup> These components can indicate the intentions of the speaker, and provides the listener with clues to properly interpret the meaning of what is being said. I have determined that the paralinguistic component is the most integral aspect of speech to be examined in this study, as it turns our attention to the relationship of form to content. For this study, non-verbal cues will be more significant to determine the intentions of the speaker (or in the case of video games, the sound designer responsible for the speech or vocal narration and cues) in communicating interactive messages to the video game player.

### 2.2.2 Music: 'Organized Sound'

Although what constitutes 'music' can be a debatable topic, I will simply be referring to music in terms of organized, composed sound for the purposes of this thesis. Music is organized sound consisting of periodic vibrations, while its antithesis would be 'noise', which, in the broadest sense of the word, refers to unmusical or unwanted sound.<sup>81</sup> It should be noted that contemporary electroacoustic music often blurs distinctions between speech, soundscape, noise and music. However, currently game designers arrange music in the traditional manner based on components of pitch, harmony and rhythm. When distinguishing music from speech, sound effects or soundscape, I am referring to composed pieces incorporating, at the very least, a melodic or rhythmic musical structure incorporating some instrumental or vocal constituent. It is

---

<sup>78</sup> Truax, Barry. Acoustic Communication. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.39

<sup>79</sup> Ibid. p.38

<sup>80</sup> Ibid. p.44-46

<sup>81</sup> Schafer, R. Murray. The Soundscape, Destiny Books, Vermont: 1994, 1977. p.182

the purpose of this section to outline the primary components that are within a musical configuration in order to understand what makes up basic music arrangement for purposes of deconstructing and isolating specific elements of music in game audio analyses. This will also aid in discerning musical elements for the sake of comprehending aesthetic and compositional decisions made by video game audio designers with regard to interactivity in video game audio. This section presumes from the reader a basic understanding of what comprises a piece of music, and in no way intends to cover the vast gamut of musical theory.

Musical styles are always changing in conjunction with shifts in musical perception by the audience. Conventional music involves notes that are generally structured and organized through a musical score, or pattern. Webster's Dictionary depicts music as an "art of sound" that expresses emotion and ideas "through the elements of rhythm, melody, harmony and color."<sup>82</sup> Music includes acoustic parameters such as loudness (shifts in intensity), pitch contours (melody), timbre (tonal quality and colouring) and rhythm (percussion, pacing). Pitch information also includes the category of chords, which are a basic part of music theory, where two or more tones are combined. Rhythm incorporates basic elements of percussion and beat structure repeated to form a recurring pattern. Music can be tightly structured and emphasized through loudness contours such as rises and falls to symbolize climactic shifts in the piece, which functions much as an acoustic envelope that can have multiple variants. A gradual increase in the overall amplitude in a climactic section of a song, for example, is called a *crescendo*, which can indicate a dramatic and dynamic shift in the song structure. This is only one example of how music can incorporate *gesture*, and can conjure up emotion in the listener as well as convey passion intended by the composer. Similar to paralinguistic considerations in speech, *how* an instrument is played in music can indicate far more than the notes themselves. Through musical structure and scoring composers have the power to create musical gestures or events that control the transitional flow and phrasing of the piece instead of focusing on the mere textual timbre of the sound at any given moment in time. Silences in the form of 'rests' can also be added to emphasize or punctuate, therein enhancing the overall meaning of a musical text. Gestural movement, as well as emphasis, punctuation, phrasing, and cadences can make musical compositions into complex artefacts that live and breathe almost independently, and can evoke emotional responses from listeners, as we will see with the introduction to psychoacoustics. Music, as we will see, can be used effectively in video game audio design to promote interactivity, as it affects the player on an emotional level. As game audio designer Aaron Marks

---

<sup>82</sup> Webster's Encyclopedic Unabridged Dictionary of the English Language. Dilithium Press, Portland House, NY: 1989. "Music" p.943

says in *The Complete Guide to Game Audio*, “Music is a powerful force in the modern world. The right music coupled with the right images is pure magic. Film and television has much to do with our perceptions, as is the music we choose for our own life’s soundtrack.”<sup>83</sup>

### 2.2.3 Soundscape: Keeping an Ear to the Ground

The term *soundscape* is adapted from the visual term ‘landscape’ to mean a sonic environment, or auditory landscape. As coined by soundscape pioneer R. Murray Schafer, any sonic environment or portion of it can be regarded ‘the soundscape’ as a field for study.<sup>84</sup> The aim of this section is to familiarize the reader with the primary auditory environmental aspects to be considered when analyzing a soundscape. As with the previous two sections, I will not attempt to cover the entire range of soundscape studies, but will illuminate the key terms to be explored for the purposes of my study. Barry Truax, who worked with Schafer on their international acoustic study entitled *The World Soundscape Project*, describes the term ‘soundscape’ as:

An environment of sound (or sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society. It thus depends on the relationship between the individual and any such environment.<sup>85</sup>

The term may refer to actual environments, or to abstract constructions such as virtual video game audio environments. For this study we will be focusing on the latter, which will incorporate sound effects, amongst other elements, within the video game audio construct (as they are related more directly to Soundscape than to either of the other categories of Speech or Music). Let us now review some of the components that make up the soundscape.

On a basic level, sounds within a soundscape can be divided into three categories: signals, keynotes, and ambience (‘soundmarks’ can be considered as a fourth category, which will be introduced later in this section). *Signals* refer to sounds that are clearly distinguished; that “stand out” to the listener. *Keynotes* are a more or less constant grouping of sounds that form a “ground” against which other signals are perceived. Keynotes are not to be confused with *ambience*, which is a similar occurrence in that it is also composed of several sounds to form a sonic environmental background. However, the key difference is that keynotes are a group of sounds that can be discerned separately from one another but stay in the perceptual background, while ambience is made up of many small sounds that are all heard together as an indivisible whole (Truax, 1984). In applying a visual landscape composition analogy, the signal would be the ‘figure’ situated atop the ‘ground’, which would be the keynotes, and all of this would be

---

<sup>83</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.187

<sup>84</sup> Schafer, R. Murray. *The Soundscape*, Destiny Books, Vermont: 1994, 1977. p.274

<sup>85</sup> Truax, Barry. “*Soundscape*”, *Handbook of Acoustic Ecology* CD-ROM, Cambridge Street Publishing, 1999.

superimposed against a 'backdrop' of blue sky, clouds and mountains, which would be the ambience in this scenario. Ambient sound in video games, as explained by game developer Daniel Bernstein, refers to the 'sound world' that exists and encompasses the player in the 'game space'. Bernstein describes sonic ambience as a "complex web of sound" that can communicate important information crucial to the understanding of the gaming experience.<sup>86</sup> According to Schafer, the term 'field' can also be added to the 'figure-ground' analogy to represent the soundscape as the place where all of the sounds occur, therein providing a framework for organizing experience (Schafer, 1977, 1994). If all of these sounds can be heard clearly to achieve an auditory balance, and sounds do not become cluttered within the soundscape, it can be termed a *Hi-Fi* soundscape. Hi-fi refers to high fidelity, and implies a favourable *signal-to-noise ratio* "in which discrete sounds can be heard clearly because of the low ambient noise level."<sup>87</sup> Conversely, a *Lo-Fi* soundscape would be one that is made up of closely packed sounds that contribute to a 'noisy' environment, disallowing the ability to discern sounds clearly from one another, thereby distorting perception. Video game audio designer Aaron Marks denotes an example of poor game sound as a "kind of mess where so much is happening at once, you have no idea what's going on."<sup>88</sup> A lo-fi sonic environment implies an unfavourable signal-to-noise ratio, which results from *masking*, which is "the effect one sound has on another by making it harder or impossible to hear."<sup>89</sup> Furthermore, Marks elaborates on the potential catastrophic consequences of an unbalanced game soundscape in the following excerpt.

An assault of unorganized audio is nothing but noise – pure and simple. With an onslaught of narratives, music, and sound effects, a game can quickly become a chaotic nightmare, forcing the player to run away screaming or to simply turn off the racket with a flip of a switch, neither of which are good for the games business.<sup>90</sup>

Achieving this balance in the acoustic design of a video game soundscape is the ultimate challenge to designers, and can make all the difference between a good or bad overall sound design. No matter how high quality the sound effects may be, nor how well composed or orchestrated a piece of game music, and regardless of how articulate and well performed a speech narrative may have been recorded for a game, a poorly designed electroacoustic environment will tarnish all of these elements.

---

<sup>86</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997. [http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm)

<sup>87</sup> Schafer, R. Murray. *The Soundscape*, Destiny Books, Vermont: 1994, 1977. p.43

<sup>88</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.261

<sup>89</sup> Truax, Barry. "*Mask, Masking*", Handbook for Acoustic Ecology CD-ROM, Cambridge Street Publishing, 1999.

<sup>90</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.261

Another soundscape term *soundmark* is derived from ‘landmark’ to refer to a unique sound that holds special information to those residing in that *acoustic community*.<sup>91</sup> This becomes an important consideration in virtual gaming environments, as soundmarks can become aural cues that symbolize something specific to the player, but which have cultural associations that extend beyond the game. Avid players of a specific game around the world (regardless of their geographic location) can be seen as united within an electroacoustic community, in that players of the game understand and identify with the same sounds and soundscape created for the gaming environment. In the study of acoustics, an ‘acoustic community’ is defined as “any soundscape in which acoustic information plays a pervasive role in the lives of the inhabitants...it is any system within which acoustic information is exchanged.”<sup>92</sup> I feel this term has relevance especially to video game players as linked together by the technology of the World Wide Web. Players of the same game, regardless of culture or geographic boundaries, can all be united within an electroacoustic community created by the fantastic, surrogate audio environment of the game. They are all linked by their understanding of an artificial soundscape, and each recognizes and identifies with the soundmarks and sonic signifiers pertinent to that gaming world.

The basic soundscape model put forth by acoustic ecologists Barry Truax and Hildegard Westerkamp serves as a foundation for understanding how sound creates a relationship between people and their habitual environment. The communicational model of how sound mediates the affiliation between listener and environment can be seen below:



Figure 5: Communicational Soundscape Model

Truax goes on to discuss the importance of context within soundscape composition, which is an integral part in the perception of any sonic environment, and certainly applies to the design of interactive audio environments in video games. Truax suggests that a sound’s meaning and context are inseparable:

... the soundscape composition is context embedded, and even though it may incorporate seemingly abstract material from time to time, the piece never loses sight of what it is ‘about’... it takes on a specific role, that of providing contextual meaning to the sounds heard.”<sup>94</sup>

<sup>91</sup> Schafer, R. Murray. *The Soundscape*, Destiny Books, Vermont: 1994, 1977. p.274

<sup>92</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed. p.66

<sup>93</sup> Truax, Barry. "The Aesthetics of Computer Music: A Questionable Concept Reconsidered." *Organised Sound*, 2000. p.122

<sup>94</sup> *Ibid.* p.124



Game developer Daniel Bernstein applies a similar perspective to video game audio, specifically with respect to ambient sound, as providing the player with information integral to an orientation with the context of the gaming environment. He states “A game with little or no ambient sound presents little or no connection to how we perceive the outside world with our ears.”<sup>95</sup>

Directly related to soundscape theory and the idea of context lies *acoustic design*, which is a central element to this study, as it refers to the analysis of a soundscape to determine how it functions.<sup>96</sup> The acoustic or electroacoustic design is not confined to the sounds themselves, but also incorporates an analysis of how sounds interrelate, and whether they contribute to achieve a balance within the auditory environment. Acoustic design combines different applications from traditional disciplines such as architectural acoustics and audio engineering to make up criteria for organizing sound waves and audio signals to devise an unobstructed *acoustic space* for the listener to discern sounds clearly.<sup>97</sup> For example, an acoustically well-designed soundscape will be a hi-fi environment that allows for a balanced, unobstructed listening experience. How an acoustic space or place is configured will ultimately affect how sounds interact within it, and subsequently how the soundscape will be perceived and appreciated by the listener. For example, a hi-fi soundscape may become increasingly lo-fi in the event of excess *reverberation*, or reflections off the surfaces of the soundscape. Reverb is the result of reflecting sound waves that bounce off the surrounding surfaces, and are heard shortly after the original sound, thereby overlapping and obstructing the potential clarity of the source sound. Reverberation time depends on the dimensions of the acoustic space, and on the amount of absorption of sound by the adjacent surfaces.<sup>98</sup> This becomes an important element of consideration for video game audio designers in creating the parameters and dynamics of virtual environments. The amount of reverberation (i.e., the number of reflections, and how long it takes for the sound to decay into silence) will provide players with important clues about their location within the virtual soundscape, as well as making them aware of their *listener position* in relation to virtual boundaries and obstacles encompassing the fantasy realm.

Electroacoustic design for an artificial environment, especially the complex fantasy realm of most video game environments, involves quite a bit of Digital Signal Processing (DSP) to enhance sounds to make them more vivid or impactful within the overall video game soundscape.

---

<sup>95</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997. [http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm)

<sup>96</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.109

<sup>97</sup> Ibid. p.13-14

<sup>98</sup> Pellman, Samuel. *An Introduction to the Creation of Electroacoustic Music*. Wadsworth Publishing, California: 1994. p.85,87

DSP can involve simple processes (mixing, equalization (EQ), panning, etc.) or can entail highly advanced digital processing techniques (time stretching, granular synthesis, effect processing, etc.) to 'tweak' the sounds through sound modification and modulation to achieve the desired effect. Truax, commenting on methods of granular time stretching, states that sound processing within electroacoustics may preserve the identity of the original sound, but furthermore "expands its qualities... to take the reference to the source sound and allow its deeper, possibly symbolic meanings to emerge."<sup>99</sup> Electroacoustic design is especially important for video game audio designers in order to create a surrogate gaming environment in which they can maintain control over individual sounds, and allow each one to be appreciated within the whole interactive construct. Techniques of multi-tracking, layering, and sampling (amongst others) are employed within different modelling tools and audio engines to function within programmed sequences to achieve interactivity. This presents some significant challenges for designers, as we will see later in this study.

### **2.3 Psychoacoustics & Auditory Perception: 'Clairaudience'**

Up to now I have been speaking about the speech-music-soundscape continuum and the acoustic and electroacoustic elements that are at play within a given environment or context. Before concluding this chapter it is necessary to discuss some of the concepts of psychoacoustic theory and auditory perception. It is important to note that I will be approaching these areas from a communicational standpoint rather than from a psychological perspective. Although this paper approaches interactive audio from a design perspective, it is still pertinent to understand the fundamentals of auditory perception, as the ultimate goal in video game audio design is to tailor the virtual sonic environment of each game to fully immerse and engage the player in order for the game to be considered effective. My aim is to inform the reader of some of the key elements of psychoacoustics and modes of perception to make an adequate reading of the ensuing chapters possible.

The study of psychoacoustics, as mentioned earlier, combines the areas of acoustics and psychology to determine *how* sounds are perceived. This has much to do with the interpretation of a sound's meaning, which is a subjective process, including semantics and aesthetics as well as taking into account values, ideologies and personal experience (Schafer, 1994). Psychoacoustics proposes that when a sound is heard, an aural experience evokes a sensation in the brain. One of

---

<sup>99</sup> Truax, Barry. "The Aesthetics of Computer Music: A Questionable Concept Reconsidered." Organised Sound, 2000. p.122

the ways to trace the meaning of the sound, or why it evokes a specific response from an individual, is by discovering the root of the sound's significance through the listener's past experiences. By understanding our responses and emotions conjured up by sounds, we can fine-tune our ears to achieve clear hearing or "*clairaudience*," the equivalent to clairvoyance in acoustics. If a sound induces a pleasurable response it is referred to as a *sound romance*, while sounds that conjure up negative emotions are termed *sound phobias*.<sup>100</sup> These concepts extend the study of psychoacoustics (which deals strictly with auditory processes) into the psychology of sound. The idea is that the sound 'resonates' in the memory of the individual, and either becomes idealized or gets stamped as having a negative connotation so that the sound carries with it a specific, subjective meaning for each listener.<sup>101</sup> Tony Schwartz likens this phenomenon to striking a "responsive chord" that resonates with the listener long after the sound stops, and becomes stored in memory as a listening experience (Schwartz, 1973). Schafer talks about the meeting of music and environmental sounds as one of the most striking features of twentieth-century music, and that the soundscape has the capacity to shift the perception of active listening. When analyzing a piece of music "psychoacoustically" (Schafer, 1977) we are tracing back to past experience and, therein, we make a connection with the music on numerous levels. Sound romances or memories can be so powerful that vivid sensations can actually be recalled, and this affects one's response to a given song. This is one of the great merits of music, that a song can evoke an intense response from someone who has no apparent link to it, and can induce a sensation for one individual or group while having a completely different impact on another set of listeners.

An audience member brings a personal experience to a reading of a given text, and this affects the perception of the text, thereby giving the person the ability to form an opinion on what is understood from the implied meaning of the content. Reactions and responses to music are not obvious. As Alan Lomax points out, "music somehow expresses emotion... one can posit the existence of a distinctive set of emotional needs or drives that are somehow satisfied or evoked by this music."<sup>102</sup> Michael Hurd, writing in 1968, states that music is not itself emotion, but is capable of "awakening in the listener strong recollections of an emotion he already knows... the emotions that music arouses [are] *real* feelings [that] may change as fast as musical phrases do."<sup>103</sup>

---

<sup>100</sup> Schafer, R. Murray. *The Soundscape*, Destiny Books, Vermont: 1994, 1977. p.146-148

<sup>101</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001. 2<sup>nd</sup> ed. p.26

<sup>102</sup> Lomax, Alan. "Song Structure and Social Structure," *Ethnology*, Vol. 1. Columbia University, 1962; p. 425

<sup>103</sup> Hurd, Michael. *The Composer*. Oxford University Press, New York & Toronto. 1968; p. 10-11

Barry Truax feels that a psychoacoustic aural phenomenon can best be understood by using the holistic term “percept” in measuring auditory activity. He feels that by shifting our focus to include the complexity of the ‘real world’, while also considering multi-dimensional psychoacoustic concepts such as timbre and volume instead of speaking of simplified psychoacoustic parameters, we can begin to understand the nature of the musical process.<sup>104</sup> Truax speaks of the effects of the electroacoustic listening experience on the listener, claiming that sound, in essence, creates a relationship between the listener and the environment in which information is exchanged and symbolic meanings emerge.<sup>105</sup> Truax uses the technique of ‘granular time stretching’, for instance, which extends the sound over time as allowing the listener to “hear the ‘inside’ of the sound in slow motion”, therein playing on the listener’s imagination to give the work its desired effect.<sup>106</sup> Peter F. Ostwald also comments on the amazing possibilities for individual perception of sonic properties, in that two completely different sounds (e.g., the roar of thunder and the patter of raindrops) can have the same meaning for the listener in suggesting stormy weather. However, he also calls attention to science’s difficulty to explain why the same word or sound can have completely dissimilar effects depending on the context (e.g., the word “hell” spoken by a television comedian and by a Priest in a church).<sup>107</sup> This feeds into the discussion of semantic meaning for the receiver, that can vary depending on paralinguistic elements and associations based on the listener’s past experience. In video games, a player’s reactions to sounds that are encountered during game play will vary depending on events that are associated with those sounds, coupled with the context in which the sounds are presented. Sound effects that are heard in games serve a key function in providing ‘aural feedback’ to the player that simulates what one would hear if one were actually in the game.<sup>108</sup> This adds to the believability of the soundscape, intended to enhance the perceived ‘realism’ of the video game as a whole.

Truax suggests that perception is dependent on the ‘level’ of listening engagement by the listener. According to Truax there are three levels of listening: *listening-in-search*, which is listening at its most active, when the ear reaches out and listens for auditory cues; *listening-in-readiness*, which depends on associations built up over time so that sounds are readily identified through “background processing”; and finally *background listening*, which is the least active

---

<sup>104</sup> Truax, Barry. “Musical Creativity and Complexity at the Threshold of the 21<sup>st</sup> Century” from *Interface*, Vol 21. 1992, p. 35

<sup>105</sup> Truax, Barry. “The Aesthetics of Computer Music: A Questionable Concept Reconsidered.” *Organised Sound*, 2000, p. 122

<sup>106</sup> *Ibid.*, p. 122

<sup>107</sup> Ostwald, Peter F. *Communication and Social Interaction: Clinical and Therapeutic Aspects of Human Behavior*. NY, London: Grune & Stratton, c1977. p. 48

<sup>108</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.290-291

form of listening that occurs when a listener is non-attentive.<sup>109</sup> Truax distinguishes between these listening modes, claiming that perception of sounds will vary for individual listeners depending on whether the subject is consciously listening for a specific sound, scanning for familiar auditory patterns, or is simply ‘tuning out’. Truax goes on to further substantiate that environmental sound signals within a soundscape, whether foreground or background in perception, can only acquire a meaning through its context.

Whether an environmental sound has meaning or not depends entirely on its context and how it is understood. The ‘sound object’ cannot *mean* anything except itself as an aural sensation. It is the ‘sound event’ that communicates depending on our ability to interpret it.<sup>110</sup>

The ‘sound event’ that Truax alludes to is the binding of a sound and its context at a given moment in time, which becomes a common term when examining video game audio from a design perspective, as will be seen later in the thesis. In speaking about computer music, which is also a component of video game audio, Truax speaks about the larger role that context plays in any soundscape composition that provides contextual meaning to the sounds heard. In composing a virtual audio environment the listener is given a generous role in interpreting the sounds in their auditory context, and many times the designer simulates the samples for the artificial world using effects manipulation to make sounds "more psychologically ‘real’ than actuality."<sup>111</sup>

Another subjective term that is often used loosely in everyday speech is *noise*. Noise symbolizes the absence of meaningful sound, and is referred to as unwanted or unmusical sound (Schafer, 1994), which is usually unpleasant to one’s ears. Noise is subjective, because in psychoacoustics “one person’s noise can be another’s music”, that is to say nothing can be universally regarded as noise. Noise has been theorized as the antithesis to musical composition (Attali, 1977, 1999), which masks out the signals we want to hear, and contributes to a chaotic lo-fi environment. Noise is a contributing factor to the escalating problem of *noise pollution*.

The final concept I wish to address within the realm of sound studies is the topic of *sound localization*. This, on the most basic level, can refer to the listener’s ability to differentiate between a sound that is coming from the left, the right, or directly ahead because of the simple fact that we have ears on either side of the head. However, increasing complexities arise with the introduction of diffracting sounds around the surface of the head, the presence of interaural time and amplitude differentials, sound colouration resulting from reflections off the pinnae, Doppler

---

<sup>109</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.19,21

<sup>110</sup> Ibid. p.45

<sup>111</sup> Truax, Barry. "The Aesthetics of Computer Music: A Questionable Concept Reconsidered." *Organised Sound*, 2000. p.124

shifts, etc.<sup>112</sup> On a very basic level, sound is received through both ears separately, which transmits sound waves to the brain to be processed as auditory information. The loudness and timbre of a sound can suggest the distance to the sound's source. For example, an instrument such as a saxophone being played at close range will sound louder in intensity, but will also sound brighter in timbre, as higher frequencies in the sound will be more prominent in closer proximity.<sup>113</sup> The most central component of sound localization is spatialization. This is the auditory equivalent to being able to perceive visual graphics 3-dimensionally (which, incidentally, has been adapted to incorporate audio in video games, as there is now the existence of *3D Sound*). Sounds are generally heard stereophonically, that is in *stereo*, which allows sounds to be perceived through both ears. Binaural, or stereophonic localization gives the effect of being encompassed by the sounds as with the sensation one feels when listening through a pair of headphones. We experience this because each channel (both left and right) has different information, or a different form of the sound being transmitted. The sounds emitted from the separate channels "combine in the air and give the illusion of direction that the human ear naturally perceives when hearing sound."<sup>114</sup> A different percentage of each sound might be played through each loudspeaker, which gives each sound source its own location between the two loudspeakers. Through a technique called *panning*, sound designers are able to transfer audio information from the left to the right portion of the stereophonic field to enhance this binaural effect. These spatial qualities are non-existent in monaural or *mono* recordings, as there is only one channel transmitting the sonic information. Building on stereophonic localization, *quadraphonic* recordings incorporate four channels instead of two. Each channel contains different information, and each of four loudspeakers is generally placed in a separate corner of a room to get the full spatial effect. "This approach expands on stereo and gives the listener a sense of sound coming from all directions. Quadraphonic recording is also the basis for today's surround-sound systems, used with home theaters."<sup>115</sup> The possibilities for surround audio are endless, now with 8-channel or *octophonic* surround capabilities, and newly implemented 16-channel systems are now on the horizon. This enhances the capabilities for video game audio designers to create an audio environment that can be fully appreciated for its spatial features, granted the technology is accessible to game players. Naturally, no game designer can truly know what percentage of players will be listening to the soundtrack that transmits its full auditory potential. However, there is always the desire to strive for the best, because "just as real ice cream

---

<sup>112</sup> Pellman, Samuel. An Introduction to the Creation of Electroacoustic Music. Wadsworth Publishing, California: 1994. p.87-90

<sup>113</sup> Ibid. p.89

<sup>114</sup> Microsoft Corporation, "Sound Recording and Reproduction." *Microsoft® Encarta® Encyclopedia*, 2001.

is preferred by discriminating taste buds to the cheap synthetic kind, so is the texture and body of the sound... favored by a well-informed sense of hearing."<sup>116</sup> We can now proceed to combine what we know about video games with the areas of interactivity and sound theory in order to set up analytical categories to analyze video game audio as creating both an immersive and interactive sonic environment.

---

<sup>115</sup> Ibid.

<sup>116</sup> Pellman, Samuel. An Introduction to the Creation of Electroacoustic Music. Wadsworth Publishing, California: 1994. P.209

*Audio in today's interactive entertainment media has progressed far beyond the bleeps of early video games. An object or an environment within a game exhibits a number of complex relationships... When encountering these elements in a game environment we expect them to communicate to us through audio in subtle and different ways. Aspects of emotion, such as surprise, frustration, admiration, and fear could easily be conveyed through an enhanced and well thought-out object vocabulary... we must not overlook the role of audio in enhancing and completing that feeling of total immersion.<sup>117</sup>*

-Daniel Bernstein  
Video Game Audio Developer

### **3. Interactive Audio in Video Games: More than Meets the Ear**

In the first two chapters I have attempted to: explain the fundamental concepts essential to a basic understanding of trends in video game audio; give an overview of interactivity and audience perception theories; and introduce primary acoustic and psychoacoustic terminology. This chapter aims to bring these three areas together for the purposes of deconstructing interactivity within video game audio as applied in the textual analyses to follow. This chapter seeks to inform the reader of some technological considerations, as well as some aesthetic design decisions that go into creating fantasy worlds that serve as interactive audio playgrounds. We will also be examining audio *algorithms* on a very basic level, to understand how interactivity, realism, and the impression of real-time arbitrary sounds are incorporated within video game audio design. In addition to exploring interactive technologies and methods for integrating interactivity in game audio, this chapter will also serve as a functional overview for the analyses to follow, and will outline the aspects to be considered and the hypotheses to be deliberated.

#### **3.1 Video Game Audio: Genres**

There are literally thousands of games spanning different genres made for different home console systems available to the video game consumer. An understanding of what distinguishes game sound in some games as opposed to others is important at this stage. It is imperative that one recognize the fact that interactive game audio will not be approached or structured in the same way for all games, spanning different genres, across different consoles. It is of the utmost importance for game audio designers to consider limitations in creating boundaries and as to what they can and cannot implement in a given game with respect to the storage and memory capacity

---

<sup>117</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997. [http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm).



of the gaming console for which their game is being designed. Critics are always quick to dismiss games as having ‘poor sound’, and gamers are even quicker to toss a game aside, claiming the audio “sucks,” regardless of how much work or how many painstaking hours may have gone into the game’s audio production. Therefore the main consideration for audio designers is not to adopt a complex formula, but simply to make audio that “works” for the game it is being created for. Creating audio that works and allows for interactivity is a challenge that can be approached in a number of ways. Let us first examine the function of audio in some primary game genres before discussing the home consoles for which they are created.

### 3.1.1 First Person Shooter Games

*First-Person Shooter* (FPS) games are action games that take place from the player’s perspective, and involve shooting weapons and engaging in combat to complete objectives and to survive within 3D spaces. This genre often contains ambient melodies and atmospheric soundtracks to set a mood and provide players with aural cues to take specific actions. The gaming environments for FPS games usually involve a high degree of interactivity, as game strategies often include moving around 3D spaces and players are forced to constantly update their playing style to either avoid, attack, hide from, pursue or ambush an enemy. In turn, the audio component of FPS games has the potential to be highly interactive, which can be programmed to adapt to the player’s style of play in real-time. Such interactive soundscapes often provide mood and atmosphere through adaptive musical scores, exquisite sound effects, and either subtle or explicit speech cues (these can range from a step-by-step vocal narration in ‘training’ or ‘practice’ stages, to subtle vocal cues to indicate a player is ‘on the right track’). In his paper *Levels of Sound* written for *VGMusic.com*, author Eric Pidkameny provides a perfect example of how interactive musical cues can vary in FPS games to adapt to a player’s given style.

...the appearance of an enemy would cause the “Enemy Theme” to play for both a cautious player and an aggressive player, but the cautious player would hear the theme played in a different style than the aggressive player would – in a minor key, for example, as opposed to a major one, or with slurred notes rather than sharply articulated ones.<sup>118</sup>

To enhance the dramatic tension in FPS games, sound effects are often strategically placed and pronounced in scoring to create a high-pressure atmosphere, much like they would be in a horror movie, keeping the player on edge, and ready to react to unexpected events and the potential for enemies lurking around every corner. Interactive audio considerations for FPS games can be quite

---

<sup>118</sup> Pidkameny, Eric. "Levels of Sound." *VGMusic.com*, *Video Game Music Archive*: 2002. <http://www.vgmusic.com/information/vgpaper2.html>.

complex, with numerous interactive tools at work to maintain the desired interactive effect. Examples of games in the FPS genre include such titles as *Doom*, *Duke Nukem*, *Medal of Honor: Frontline*, *007 Nightfire*, *GTA*, *SOCOM: US Navy Seals*, etc.

### 3.1.2 Graphic Adventure Games

*Graphic Adventure* games, in contrast to FPS games, rely less on themes of shooting and killing, and instead focus players' energies on problem-solving and exploring the virtual environment that encompasses them. These adventure games adopt narrative styles similar to those of film in developing detailed characters and intricate storylines that evolve and become revealed gradually to the player as the game progresses. Gamers are often required to explore various locations in order to complete multifaceted objectives that can be quite complex, requiring the player to accomplish a number of tasks in order to progress to the next level. Graphic Adventure games have evolved from *text adventures*, one of the earliest computer game genres. Text adventures in the 1980s, such as *Zork*, *Planetfall*, and *Space Quest*, used text rather than graphics to describe the characters and gaming environment because computers lacked the capacity to store graphic image or sound files, and required players to enter commands on a keyboard.<sup>119</sup> In contrast to the text-based adventure games of twenty years ago, their predecessors, graphic adventure games, are rich and sophisticated, both graphically and sonically. Sound for graphic adventure games tend to be more subtle and ambient. The goal is for music and sound to create a mood in a non-intrusive manner to guide the player through an engaging soundtrack, and paying special attention not to distract the player. It is in these games that music composers have garnered some much deserved attention in recent years for gestural and rich orchestral compositions. The challenge for game audio designers and composers in these games also involves creating an interactive score that has many subtle shifts so as not to allow the soundtrack to become overly redundant or repetitive, seeing as how audio sections may be looped several times over while players attempt to solve the game's puzzles. Audio interactivity is a struggle in graphic adventures because gamestates can last anywhere from a few minutes to several hours, depending on a player's actions. Some of the more popular graphic adventure games include *King's Quest*, *Loom*, and the *Myst* series.

---

<sup>119</sup> Ibid.

### 3.1.3 Role-Playing Games

*Role-Playing Games* (RPG) have a similar style to the aforementioned graphic adventure games in placing emphasis on exploration and development of the storyline. However, RPGs differ because the main focus is on fending off monsters and protecting the fantasy world from evil rather than solving puzzles. These games focus quite a bit on character development as well, cultivating a character's persona through traits, race, gender, strength, agility, endurance, etc. Because there is more emphasis on building intricate characters and a nuanced, composite plotline in RPGs than in other game genres, there is also far more potential for an extensive soundtrack. Sadly, however, these games are least likely to incorporate sonic interactivity, as they are mostly dependent on a pre-set narrative structure that allows for little deviation from the pre-programmed storyline and auditory accompaniment.<sup>120</sup> It is also important to note that numerous battles occur at different stages of the game, therefore the soundtrack accompanying the battle event will usually be heard more often than any other audio sequence in the game. If poorly composed or arranged, this can be detrimental to the overall sound design of the RPG, as it can become repetitive and annoying to the player. Although the music accompanying the battle gamestate can be riveting, varying according to level, difficulty, health, rate of success, speed, etc., these could also become repetitive and redundant for the player, creating an eventual detachment from the game. Initially players would identify with the sequence, thus creating a sense of familiarity, although the drawback to this cyclical nature of the same audio event repeating over and over again would eventually limit the 'replay value' (how often the player resumes play of a given game). Even the player's favourite song would grate on one's nerves when hearing it repeated hundreds of times in a long day of gaming.<sup>121</sup> The end result is often the eventuality that the player may become frustrated and could discard the game, or worse, the user may turn the sound off altogether (a sound designer's worst nightmare). By the same token, the soundtrack of an RPG, especially with respect to its musical score, can be beautifully composed and orchestrated, much like melodious, dramatic film music. Composer Nubuo Uematsu of the acclaimed *Final Fantasy* series has been known to orchestrate rich, sophisticated musical scores in his compositions, as has Michael Giacchino, composer for *Medal of Honor: Frontline*, a 1999 game inspired by the Spielberg blockbuster film *Saving Private Ryan*.<sup>122</sup> Giacchino's compositions feature full orchestral scoring, and the lush, complex movements have been likened

---

<sup>120</sup> Ibid.

<sup>121</sup> Harland, Kurt. "Composing for Interactive Music." 2000.  
[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

<sup>122</sup> Coleman, Christopher. "A Conversation with Composer Michael Giacchino." 2001.  
<http://www.tracksounds.com/specialfeatures/Interviews/interviewing.htm>

to the award-winning compositions of film soundtrack guru, John Williams.<sup>123</sup> Prime examples of RPGs include *Dungeons & Dragons*, *Diablo*, and the ever-popular *Final Fantasy* series.

### 3.1.4 Sports Games

*Sports* game titles generally focus on immersing the player into a competitive gaming environment tailored to fit the style of the game. This is to say the audio can either take on realistic connotations, or can be fantastic and over-the-top, depending on the desired and goals of specific programmers. The degree of realism will guide the audio process and overall design structure.

Sports game simulators (sims), fashioned after actual, real-world, professional leagues, will go for as much realism as possible to immerse the player in a realistic gaming experience. Long past are the days of computer-generated voices with vocabulary limited to “strike”, “ball”, and “out”, as were the speech characteristics of the first talking sports game *Major League Baseball* for the *Intellivision* home console system in 1979.<sup>124</sup> Audio designers are currently aware that avid players of such games have come to expect a gaming experience that is as true to a real-life sports context as possible. The more realistic the game experience, the more successful the game will be. Statistics and game scenarios are studied and implemented in the games so that all game elements will interact the way they would in the real world. Any deviation from what would constitute a ‘real’ sporting event will detract from the quality of gameplay. For example, graphically, games like *NHL 2003*, *NBA 2K3*, *Triple Play Baseball 2002* and *FIFA Soccer 2002* will all attempt to design players and stadiums that fit the specifications of the actual real-life figures they are fashioned after. The architectural 3-dimensional construct of the stadiums are designed to look like the originals, and the players are virtually identical to the people they portray to details of appearance, mannerisms, strength, agility, endurance, etc. Audio functions in much the same way, in that the audio attempts to be as realistic a representation as possible, incorporating the actual voices of the players, tailoring crowd ambience to reflect the geographical context of where games are being played, the music that the virtual stadium’s PA system will play during different game events or player introductions, etc. Play-by-play commentary is also a driving force in such games, which attempt to convey a realistic narration of the game experience, identical to the random commentary that would accompany a professional sports telecast. Technology and speech tools have evolved so drastically in recent years that the

---

<sup>123</sup> Payne, Joseph. "FilmTracks Modern Soundtrack Reviews: Medal of Honor." 2001.  
[www.filmtracks.com/titles/medal\\_honor.html](http://www.filmtracks.com/titles/medal_honor.html)

<sup>124</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music."  
[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html)

dialogue between the play-by-play announcer and the color-commentator sounds completely believable, corresponding to game events as they happen in real-time, and flowing as human speech would in a real conversation. This is a far cry from the primitive play-by-play commentary first introduced in Sega's *Joe Montana Sportstalk Football II*, and taken to the next level with each generation of the *Madden Football* franchise<sup>125</sup> which, although they featured the first continuous commentary describing game events, sounded extremely robotic and inhuman. As we will see in the case study, ambient crowd soundscapes coupled with music, speech and sound effect overlays can be extremely complex to simulate a real-time crowd that reacts to every scenario or gamestate that can be executed by the player in interactive audio.

Conversely, fantasy sports games will feature exaggerated moves, actions, scenarios, and events to supersede what could be attempted or achieved in the real world. Games that portray extreme sports often gravitate to this genre. Players of these games are said to expect over-the-top graphics and sound, and any exaggeration of realistic parameters would be presumed as welcomed rather than detracting from the gaming experience. This is not to say that the games are *unrealistic*. Fantasy sports games will still feature realistic moves and events to the actual sports they emulate, but the graphic and sonic simulations often strive to have a greater impact than they would in real-life to glorify the gaming experience. The voices of the characters will be more animated, crowd cheers and heckles will be exaggerated, and speech commentary can range from the comical to the absurd, depending on the target age group for the game. Interactive music for such games can often take on a greater role than in realistic sports simulation games, as they can be incorporated as foreground elements to guide gameplay, rather than as background ambience as in realistic stadium soundscapes. Music, while used sparingly in sports simulators, can be incorporated to adapt to gameplay, while simultaneously encouraging the player through reward music, all the while providing an adrenalizing mood through an upbeat, fast-paced soundtrack. Some fantasy sports games include *SSX Tricky*, *FreekStyle*, *SledStorm* and *Tony Hawk's Pro-Skater 3*.

### 3.1.5 Music Simulation Games

The last game genre I'd like to mention is what I will call the *Music Simulation* game. These games are relatively new, building on the wide array of video game genres that is constantly expanding. Introduced in SCEA's revolutionary 1997 game *PaRappa the Rapper*, a new gaming experience was born, combining the distinctive practices of music-making and

---

<sup>125</sup> Ibid.

techniques of video gameplaying. Such games serve as musical interfaces that allow gamers of all skill levels, requiring no prior knowledge of musical techniques or practice, to create a soundtrack using the buttons and joysticks of home console game controllers as instruments. By pushing the appropriate buttons in time with the accompanying hip-hop beats of the game soundtrack, the goal is to get points by adapting the soundtrack to the game's specifications, as the combinations get more difficult, and the tempo gradually increases. The premise was taken a step further with Konami's arcade game *Dance Dance Revolution* in 1998, which required the player to step on dance pads in time with the music to earn points. The game is just one of many "Benami" music arcade games to soar in Japan. Others include *Guitar Freaks* (in which the player must play a toy guitar to music), *DrumMania* (the object of the game being to play on a drum kit peripheral device in time with the music), and *Hip Hop Mania* (which allows players to hone their deejay skills in scratching turntables to musical accompaniment).<sup>126</sup> The home console market has latched on to this phenomenon with *Frequency* for the Playstation 2. The game has a similar premise to all of the above music simulation games, but contains a twist. *Frequency* contains a 'remix' function that allows the player to interact with the song, in rearranging and structuring the original instruments and musical elements of the song to their own specifications, therein creating entirely new musical pieces that can be saved on their system using a memory card for storage. The level of interactivity provided by music simulation games is debatable. The fact that users are in essence creating new musical arrangements in time with computer-generated beats allows these games to be seen as having interactive potential. However, it can also be argued that interactivity is limited because the object of the game usually requires players to follow a pre-determined song structure, and to press buttons when they are instructed rather than devising their own path within the gaming environment.

### 3.2 Video Game Audio: Home Console Systems

Without delving too deeply into the technical parameters of home console technology, it is necessary to mention the three primary home console systems as vehicles for which the above genres of games are created. The relatively simultaneous launch of the Sony PlayStation 2, the Nintendo GameCube, and the Microsoft Xbox constitute the new era for video game audio development. Unlike previous methods for implementing video game audio, sound artists are now able to use familiar tools instead of having to deal with the often painstaking task of learning whole new audio development systems specific to each console.<sup>127</sup> It is important to note that

---

<sup>126</sup> Ibid.

<sup>127</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.273

there are a number of different home gaming consoles available, each with different hardware and software specifications that dictate the capacity for storage. Having a basic understanding of the audio capabilities and limitations of each of the three top-selling home game systems is necessary in order to fathom how and why game audio was designed for the specific games to be analyzed in the fifth chapter.

### 3.2.1 Sony PlayStation 2

Sony's release of the *PlayStation 2* (PS2) in 2000 set a new standard for video game technology. Being the first of a new generation of home consoles for the new Millennium, Sony broke new technological ground by implementing the PS2 with a main processing unit known as the Emotion Engine (EE), a central processing unit (CPU) with the capability of 128 bits and 295MHz speed supported by 32MB of random access memory (RAM).<sup>128</sup> Sony quickly became a hot commodity, providing the ability to link to the Internet and play DVDs in addition to games. The PS2 console made significant technological advances from its predecessor, the 32-bit PlayStation, released five years earlier. In addition to quadrupling its graphical processing capabilities, the PS2 also doubled the original PlayStation's 24-channel sound chip to a 48-channel system.<sup>129</sup> The PS2 also expanded sound barriers by creating an internal sound processor that exists independent from graphics processors to keep from having to share its resources with visual graphics, which historically have garnered greater emphasis and devoted CPU memory. The sound unit is known as the Input/Output Processor (IOP) that features two complementary SPU2 chips that combine for a total of 2MB of memory devoted exclusively to sound.<sup>130</sup> The console allows for high quality sounds up to 48kHz sampling rates (digital, CD-quality sound), although designers usually still use audio files of a lower sampling rate in order to conserve space. Audio files are streamed separately within the IOP, and then mixed in the SPU2 using up to 48 total voices that can be played back at once at any given stage of a PS2 game.<sup>131</sup> The 2MB of sound allows for higher quality audio that can be stored directly in RAM, making it possible to stream continuous music, sound effects, speech components and background ambience simultaneously, while maintaining stereo or surround sound localization.<sup>132</sup> Naturally there are critics of the PS2, who claim that the advances in audio technology, although they greatly improve sound quality and storage, are still overshadowed by advances made to improve graphics support to make games that look better (Marks, 2001). However, although visual elements and

---

<sup>128</sup> Ibid., p.274

<sup>129</sup> McDonald, Glenn. "Gamespot: Brief Timeline of Game Music." 2001.  
[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html).

<sup>130</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.274

<sup>131</sup> Ibid., p.274

graphic advances take up the bulk of the console's resources, it is still significant that the PS2's independent sound processor has opened new doors for sound designers. Audio artists designing for the PS2 can implement techniques with fewer restrictions, having the opportunity to preserve high-quality audio, incorporate longer loops, larger sound samples, more processing and modelling capabilities, and 48 independent voices that can be streamed simultaneously in stereo sound. Sony's PlayStation 2 is still the most popular and most sought-after home game console of the Sony/Nintendo/Microsoft triumvirate, and is at the forefront with the highest number of games currently being produced and manufactured in the video game market.<sup>133</sup> The world leader in interactive electronic entertainment, *Electronic Arts* (EA), who currently hold approximately 15% of the video game market and posted revenues of more than \$1.7 billion for fiscal 2002, develops the majority of its games for the PS2.<sup>134</sup> Because of the factors I have just mentioned I felt it would be appropriate to focus on games for the PS2 home console when choosing specific games to analyze for this study.

### 3.2.2 Nintendo GameCube

In creating the *GameCube*, Nintendo converted its format to optimal CDs instead of its previously limiting game cartridges, which allows for larger file sizes for both graphics and audio. The GameCube's main controlling unit is "a 405MHz, IBM Power PC, 3D enhanced microprocessor unit known as 'Gekko'."<sup>135</sup> System memory and audio data are stored in 24 MB of IT-SRAM, which is considered to be the fastest RAM currently available.<sup>136</sup> Similar to the PS2, the GameCube is also capable of 48kHz quality stereo or surround sound, but has a digital signal processor (DSP) that supports more than 100 voices, and can play up to 64 real-time 3D voices at one time.<sup>137</sup> The beauty of the GameCube, from a sound composer designer's standpoint, is that there is no need for a development kit to create audio for Nintendo games. Music and sound effects can be produced in whatever method the sound artists see fit, and can be implemented and easily inserted into the game by the programmer using Nintendo's audio tool called *MusyX*.<sup>138</sup> The MusyX tool handles all of the audio needs for a GameCube title, including

---

<sup>132</sup> Ibid., p.275

<sup>133</sup> Pidkameny, Eric. "Levels of Sound." VGMusic.com, *Video Game Music Archive*: 2002. <http://www.vgmusic.com/information/vgpaper2.html>.

<sup>134</sup> Electronic Arts Inc. "Investor Info." 2002 [http://www.corporate-ir.net/ireye/ir\\_site.zhtml?ticker=ERTS&script=2100](http://www.corporate-ir.net/ireye/ir_site.zhtml?ticker=ERTS&script=2100); *About EA*, 2002 <http://www.ea.com/global/corporate/aboutea.jsp?src=11001homecorpabounone>.

<sup>135</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.276

<sup>136</sup> Ibid., p.276

<sup>137</sup> Satterfield, Shane. "Gamespot Presents: GameCube Dossier" *Gamespot*. 2002.

[http://gamespot.com/gamespot/features/video/gamecube\\_dossier/p2.html](http://gamespot.com/gamespot/features/video/gamecube_dossier/p2.html)

<sup>138</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.276



sound effects, streamed audio and interactive music sequences. 3D surround functions such as amplitude, sound envelope, frequency, panning, and effects processing can also be created using this tool, and can be “randomized” without relying on loop points as references, but rather through banks of sounds that are programmed as “sound events.”<sup>139</sup> In other words, audio designers developing for GameCube titles can create and apply components of different sounds separately, and then have them played back at random, which can then be layered according to events, ultimately allowing for more interactivity.

### 3.2.3 Microsoft Xbox

Last, but certainly not least, comes Microsoft’s contribution to the home gaming console stage. Microsoft lives up to its reputation as a computer hardware and software giant with the *Xbox*, which has capabilities that exceed both Sony and Nintendo, with an 8GB external hard drive and Ethernet port supporting DSL and cable broadband Internet connections.<sup>140</sup> Some of the *Xbox*’s features include a Pentium III processor with 733mHz CPU, 64MB RAM, and DVD capabilities. The *Xbox*’s audio data unit is the Media Communications Processor (MCP), which includes four independent audio processors for digital audio and 3D positioning; the *set-up engine*, *Voice Processor (VP)*, *Global Processor (GP)*, and *Encode Processor (EP)*.<sup>141</sup> The processors each perform a specific function to ensure high-quality digital audio. Unique to Microsoft is *DirectMusic Producer*, a composition tool used in conjunction with *DirectX 8* audio scripting, *Downloadable Sound (DLS)* and *Windows Media* software for creating and programming game sound.<sup>142</sup> *DirectMusic Producer* is a program that can be used to generate interactive audio, as will be explained later in this chapter. The *Xbox* has a definite edge in the audio department in that it will eventually be the first console with the ability to incorporate in-game Dolby Digital surround sound, as well as the capability of communicating 256 simultaneous stereo voices through 64 discrete channels.<sup>143</sup> This should solidify the *Xbox* as having the potential for creating a far superior auditory experience to that of the PlayStation 2 or GameCube.

---

<sup>139</sup> Ibid., p.277

<sup>140</sup> Gamespot.com. “Gamespot Presents: Xbox Dossier” *Gamespot*. 2002.  
[http://gamespot.com/gamespot/features/video/x\\_box\\_dossierép1\\_01.html](http://gamespot.com/gamespot/features/video/x_box_dossierép1_01.html)

<sup>141</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.279

<sup>142</sup> Ibid., p.279

<sup>143</sup> “Gamespot Presents: Xbox Dossier” *Gamespot*. 2002.  
[http://gamespot.com/gamespot/features/video/x\\_box\\_dossierép1\\_01.html](http://gamespot.com/gamespot/features/video/x_box_dossierép1_01.html)

### 3.3 Technological Advances in Video Game Audio: Loud & Clear

Just over three decades have elapsed since Allan Alcorn designed the historic PONG video game for home use on the early Atari console system. The game was so-named for its audio component, a hollow ‘bleep’ noise that sounded each time the ball made contact with an object on the screen. Although President of Atari Inc. Nolan Bushnell had high aspirations for integrating game sound, Allan Alcorn himself admits that sound for PONG was nothing more than a lucky accident:

People have talked about the sound, and I’ve seen articles written about how intelligently the sound was done and how appropriate the sound was. The truth is, I was running out of parts on the board...I said, “Screw it, I don’t know how...I don’t have enough parts anyhow.” Since I had the wire wrapped around the scope, I poked around the sync generator to find an appropriate frequency or a tone. So those sounds were done in half a day. They were the sounds that were already in the machine.<sup>144</sup>

Regardless of whether the sonic blips heard in PONG were stumbled upon accidentally or not, they still serve as the first instances of video game audio, and set the stage for the revolutionary, artistic, and powerful soundtracks that exist today. It can be said that advances in video game audio technology have followed a similar trend to that of its close media cousin, the motion picture. Much as the film industry experienced when sound was first introduced into movies, the aural experience in video games has evolved from electronic bleeps and synthesizer drones, and is gravitating toward the vivid and dramatic scores found in today’s movies. However, as we will see in the remainder of this chapter, factors of interactivity are what distinguish video game audio from film scoring. The rest of this chapter will examine the evolution of video game audio in recent years from fixed, linear methods of stationary audio (much like that of a film soundtrack), into the adaptive, event-driven, immersive and adrenalizing musical and environmental soundtracks they have become.

#### 3.3.1 Linear Audio

Linear presentation implies audio that flows from beginning to end in a predetermined and scripted way. An audio soundtrack is composed and arranged by the composers and designers beforehand, and is entirely programmed to play back in a chronological fashion. In this type of audio construct, a player would be positioned much as a film audience member; a viewer passively watching and listening to the plot unfold without having any ability to alter the sequence of events, or to interact with the storyline. Of course it can safely be argued that the

---

<sup>144</sup> Kent, Steven L. The Ultimate History of Video Games. Prima, Random House: New York, 2001. p.42

audience member still engages with the story on an emotional level, and that the soundtrack, especially with respect to the musical scoring component, has the potential to immerse the listener within a powerful, complex and moving score. As Aaron Marks states “Music is a powerful force in the modern world. The right music coupled with the right images is pure magic. Film and television has as much to do with our perceptions, as is the music we choose as our own life’s soundtrack.”<sup>145</sup> There is no question that film sound has the power to engage the audience, and provides emotional stimuli to evoke responses from the audience. Movies are an excellent way to examine how dramatic musical scoring can be used to stimulate emotion in the audience, enhancing what characters are feeling in a way that independent visuals cannot. In this sense, video game music is much like film scoring, having similar purposes in “creating a mood, setting the pace, highlighting plot shifts, and adding tension and excitement to all of the appropriate spots.”<sup>146</sup> On an emotional level, Daniel Bernstein says the soundtrack for video games “carries the person seamlessly along with the action in much the same way as the static, linear medium of film.”<sup>147</sup> However, for the purposes of distinguishing between interactive audio considerations for video games and the linear audio structure adopted by motion picture scoring, it is important to understand that the potential for interactivity is simply not made available with regard to a linear audio construct in which the soundtrack begins, plays, and then finishes. Denis Labrecque, in his article *Altering Algorithms to Create '3D' Sound*, pinpoints linear sound as “inefficient” in attempting to preserve realistic representations in video game audio with “simple static-wave playback and repetitive loops.”<sup>148</sup> Most video games cannot adopt this method, as the goal of video games is for something different to happen each time the game is played, and for the audio to adapt and flow along with the gaming experience as determined in real-time by the player’s actions. Not all games necessitate an adaptive musical accompaniment more than relying on game audio to set a mood, as outlined in the earlier *Video Game Audio: Genres* section. Currently this non-linearity and reactive, adaptive nature strived for in game audio presents significant challenges to game audio designers in attempting to create a soundtrack that evolves with the storyline as determined by the player. Renowned video game audio composer and designer, George Alistair Sanger, a.k.a. The Fat Man, compares audio interactivity to a ‘weapon’ in the following passage.

Audio, especially game audio, is a powerful weapon. When used properly,

---

<sup>145</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.187

<sup>146</sup> Ibid., p.189

<sup>147</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997. [http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm).

<sup>148</sup> Labrecque, Denis. "Altering Algorithms to Create '3D' Sound." *EE Times*. 2001. <http://www.etimes.com/story/OEG20011218S0047>.

it has the power to involve, immerse, elevate, and reward. It has the power to excite. It can make an artificial world appear to be deeper, older, and much more complex and complete than it actually is. But when misused, audio reveals its most awesome and deadly power – the power to annoy.<sup>149</sup>

Sanger warns about the dangers of attempting to implement linearity in video game audio design, using analogies of a game developer's one hour of music being stretched over a 40-hour entertainment experience as being likened to a box of crayons filled with only one colour, or "driving cross-country with one audio cassette that you didn't choose."<sup>150</sup> The Fat Man pinpoints repetition in game audio as the enemy, and warns that previously adopted methods of incorporating conventional music or film scoring techniques to video games don't allow for variety or audience feedback which is what makes interactive entertainment unique. Political Economist Jacques Attali devotes an entire section to *Repeating* in his book *Noise*, in which he pinpoints the repetitive character of music as a form of social control, contributing to the "death of the individual", the "stockpiling death" of music, and the "cancer" that is killing our freedom to make meanings and interpretations.<sup>151</sup> This is further presented in the theories of acoustic ecologist Hildegard Westerkamp, who describes Muzak as contributing to a musical culture that is "devoid of meaning," thus creating a commercial world of commodity exchange in which the listener becomes nothing more than the "consenting purchaser."<sup>152</sup>

For early games with sound, film-scoring techniques were a distant goal that could not realistically be met for video game audio designers. The goal was to evolve beyond stationary, mono, 8-bit electronic sounds that became typical as the cheap video game sounds players became accustomed to. Game music incorporated basic synthesizer technology, which introduced new sounds, textures and methods for manipulating sounds. Audio for video games wasn't generally concerned with more than basic techniques of stimulus-response, cause-effect relationships of sounds within the artificial environment. Game developers and programmers were more concerned with devoting financial resources to improving the look of the game, to make the visuals more attractive. Sound artists were not even employed at this stage, and instead graphical engineers attempted to implement whatever audio they could manage to add some form of auditory representation of the visual components. Budgetary restraints and technological limitations always became an issue in the design process, and audio was often left as an afterthought. Game programmers were not concerned with real-time computing with respect to

---

<sup>149</sup> Sanger, George Alistair. "The Sound of Money (Down the Potty): Common Audio Mistakes in Kids' Games." Gamasutra audio archives, 2001: [http://www.gamasutra.com/features/20010515/sanger\\_01.htm](http://www.gamasutra.com/features/20010515/sanger_01.htm).

<sup>150</sup> Ibid.

<sup>151</sup> Attali, Jacques. *Noise: The Political Economy of Music*. The University of Minnesota Press, 1985. p.127, 132

audio, but would often demand simply what the game needed in order to be remotely believable in the context of the gaming environment, in other words sounds that would reflect or represent the images, content to picture. The premise was that the audio must fit the visuals, and the aural effect must be designed and synchronized with the visuals. Music for games was created to fill the void of silence, and to create a mood for gameplay. Sound effects were created based on animated elements.<sup>153</sup> Game audio consisted of music and ambient tracks coupled with voice and sound effect samples that often made for a very static and fixed, unchanging environment. One of the earliest sound games, *Space Invaders*, featured the first real instance of ‘non-diegetic sound,’ which is a term used in film that refers to sonic elements “amplifying the mood of the scene and/or explicating dramatic developments and aspects of character.”<sup>154</sup> The music in *Space Invaders* increases in rhythm to create intensity when the player does not destroy the enemy alien characters quickly enough. Of course the sound quality improved over time, but the engagement between the gamer and the virtual environment was limited by the synthetic quality of a fixed story line, and static characters, all of which were represented by electronically synthesized and robotic sounding audio. Music sounded synthesized, speech sounded robotic and the soundscape conveyed a static, stationary environment that gave no indication of listener position or binaural localization.

Until more recently most current titles made use of standard linear music and sound effects that were designed to play from start to finish (Marks, 2001). Because technology hadn’t evolved to include techniques of implementing interactivity through audio, the focus was to include audio of superior quality, and to create the appropriate atmosphere and mood to create tension, giving the sense that something is about to occur as in film scores. The introduction of *Musical Instrument Digital Interfaces* (MIDI) in 1982 gave game audio designers and composers a new playing field, and adopted a significant role in game music because it was a versatile format that could produce synthesized music in files of relatively small sizes (Marks, 2001). The establishment of the MIDI protocol also made it possible for electronic instruments to communicate with each other, and with computers, as MIDI files simply contain data specifying when and how long to play each note. MIDI, although it opened new doors for game audio, was extremely limited by its lack of support for audio interactivity, and the fact that MIDI had timbral inconsistencies when played back on different sound cards, detracting even more from auditory

---

<sup>152</sup> Westerkamp, Hildegard. “Listening and Soundmaking: A Study of Music-as-Environment” from *Sound By Artists* (Lander & Lexier eds.) 1990. p.228

<sup>153</sup> Clark, Andrew. “Designing Interactive Audio Content-To-Picture.” Gamasutra, 1999. [http://www.gamasutra.com/features/19991220/clark\\_01.htm](http://www.gamasutra.com/features/19991220/clark_01.htm).

<sup>154</sup> Macy, Laura. Ed. “Film Music.” *Groove Music*. *The New Grove Dictionary of Music and Musicians*, 2002. <http://www.groovemusic.com/index.html>.

realism.<sup>155</sup> By the mid-80s video game music that ‘sounded’ like conventional music became possible, and producers insisted on having song-like structures applied to game audio. Hiring game composers and sound designers to implement music that was enjoyable to listen to became a priority in game audio, and musical gaming trends evolved toward implementing pop music styles into games. Unfortunately this was yet another setback for audio interactivity, in that music for listening requires composition that is linear, fixed and invariable, thus imposing limitations.<sup>156</sup> The emphasis on composing linear music for the sake of listening detracted from techniques of interactivity. The problem was that compositional approaches to creating a well-arranged song are methods that must be conceptualized differently for interactive composing. As interactive video game music composer Kurt Harland writes “Composing an interactive piece requires that we imagine our song sounding any number of different ways, and this is something we’re not used to.”<sup>157</sup> Sampling technology became available in the early 1990s, which gave composers and designers the ability to create rich layers of digital electronic instruments and acoustic sources.<sup>158</sup> At this stage audio interactivity in video games was almost forgotten.

Computer music authoring systems and tools were made available for composing game music in the 90’s with emphasis on sound quality. The three most widely used authoring systems were *General MIDI*, which made use of built-in synthesizers and sound cards, *Redbook audio*, which makes use of digital CD-quality audio, and *MODs*, that are digital modules that play built-in samples through a digital chipset.<sup>159</sup> All of the above methods attempt to make game composing simpler, and improve the quality of game audio. MIDI was an easy solution, in that it took up little memory, and made synthesizer music easy to manipulate for game scoring. Redbook audio eliminated the problem of poor quality game audio, allowing music to be streamed directly into the game, and giving game audio quality the movie soundtrack quality it had long been striving for. With MODs composers had even fewer creative boundaries in designing and streaming digital music and sound through sequencing programs. However, each of the methods began to be criticized for their individual flaws. Game developers began to despise MIDI for its unrealistic and inconsistent sounds that vary depending on the sound card being used by a specific system. MODs are highly complex and difficult to master, with all of the parameters

<sup>155</sup> Yackley, David. "Microsoft DirectMusic: Creating New Musical Possibilities." Microsoft Corporation, 1999. [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dm/music/html/dm\\_nmp.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dm/music/html/dm_nmp.asp).

<sup>156</sup> Harland, Kurt. "Composing for Interactive Music." Gamasutra, 2000. [http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

<sup>157</sup> Ibid.  
<sup>158</sup> Yackley, David. "Microsoft DirectMusic: Creating New Musical Possibilities." Microsoft Corporation, 1999. [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dm/music/html/dm\\_nmp.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dm/music/html/dm_nmp.asp).

<sup>159</sup> Brandon, Alexander. "Interactive Music: Merging Quality with Effectiveness." Gamasutra, 1998. [http://www.gamasutra.com/features/sound\\_and\\_music/19980327/interactive\\_music\\_quality\\_intro.htm](http://www.gamasutra.com/features/sound_and_music/19980327/interactive_music_quality_intro.htm).

needing to be programmed and defined in order for it to sound as the composer intended it. Redbook audio, although maintaining high quality audio, requires a substantial amount of memory and processor power to play, which competes with built-in RAM devoted to graphics and limits interaction between the soundtrack and the player.<sup>160</sup> The artificial intelligence of characters in the game are made up of complex personalities, attitudes, actions and characters traits, and furthermore, the vastly expanding trend of online gaming allows human players to face one another through game forums. However, online gaming communities using broadband connections to link players through the web have not evolved to audio's current potential and sophistication incorporated in console gaming. As has been the case with any new audio development in video games, the emphasis is on improving sound quality rather than the interactive audio experience. This will be discussed in the concluding section of this thesis on the *Future of Interactive Audio in Video Games*.

Now that CDs, DVDs and mp3 sound files have become the norm, along with 48kHz, 16-bit stereo samples within RAM devoted exclusively to sound, home console audio quality has reached its pinnacle as far as digital quality, detail, and dynamic range of sound. Although some musicians are working with sampling rates as high as 96kHz, sound quality, resolution and file sizes have reached their optimal requirements.<sup>161</sup> RAM is the element of memory exclusive to console gaming that makes audio storage different than arcade and computer gaming. It is only in the last five or so years that interactivity has become a reality for video game audio designers. Although there have been subtle advances to incorporate audio that integrate musical variation as the video game soundtrack evolves, it has only really been possible to implement audio that adapts to gamestates and player's real-time actions with recent technological advances in home consoles of the new millennium. With expanded sound memory storage allocation, as well as new tools for composing, designing and mapping interactive soundtracks that evolve and adapt through gameplay, the stage is set to tailor interactive audio to what the player will experience as the game evolves. Interactive audio capabilities, primarily with respect to music, are being slowly realized and adopted, although some passive film techniques are still integrated into interactive gaming in the form of *Cutscenes*. These are essentially mini-movie segments within the game, during which the player has no control, and simply watches (and listens) the plot evolve.<sup>162</sup> While interactivity in game audio still has not reached its full potential, the audio is well on its way to

---

<sup>160</sup> Ibid.

<sup>161</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.175

<sup>162</sup> Macy, Laura. Ed. "Film Music." *Groove Music: The New Grove Dictionary of Music and Musicians*, 2002. <http://www.grovemusic.com/index.html>.

matching the player's moves and style of play, following twists in the evolving plot specific to the player.

### 3.3.2 Interactive & Adaptive Audio

In contrast to film audio, game music now aims at providing an interactive experience exceeding that of a film soundtrack. To create an engaging, participatory sonic experience, game audio is referred to as interactive when it can be altered directly by a player's actions in real-time. Rob Ross suggests that "Interactive audio is a technology designed to allow specifically created audio, placed in a given application, to react to user input and or changes in the application environment."<sup>163</sup> Contrary to film audio techniques, Daniel Bernstein discusses the challenges inherent in creating an interactive audio environment.

When we go to a movie, our emotional response is directly related to the music. The music swells, our anticipation grows, and our adrenaline rushes. The music ebbs, and we feel a calming sensation. This is very easy to convey in a linear medium, where the ending and the progression of events in a movie is predetermined; but how do we compose a soundtrack to a game if it can follow many paths and endings? An adaptive soundtrack that responds well to game events is one of the best ways to envelop the player in a game experience.<sup>164</sup>

Furthermore, Kurt Harland states that interactive game audio composition is unlike film scoring because video game audio designers do not have the luxury of knowing what will transpire on the screen first and composing to fit the visual components; game music must be incisive and dramatic and must set the emotional tone of the character's experiences.<sup>165</sup> Liam Byrne, a sound card specialist for Creative Labs, cites film music as a model for interactive audio, although game soundtracks go beyond the realm of film scoring in providing an interactive experience.

We want to take the experience that everybody has at the movies and make it into something that you control... You're playing through your own adventure. We're used to constant soundtracks in your entertainment. The more exactly the video game soundtrack matches your experience, the more involving that experience is going to be... Entertainment is likely to become more and more immersive as time goes on, and perhaps in the future there will be games that are very much like stepping inside a movie; if gameplay is moving in that direction, then game music will somehow have to move along with it, providing soundtracks that are as involving

---

<sup>163</sup> Ross, Rob. "Interactive Music...er, Audio." Gamasutra, 2001.

[http://www.gamasutra.com/resource\\_guide/20010515/ross\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/ross_01.htm).

<sup>164</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997.

[http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm).

<sup>165</sup> Harland, Kurt. "Composing for Interactive Music." Gamasutra, 2000.

[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)



as a film score but as specifically personalized as a fingerprint.<sup>166</sup>

However, Rob Ross maintains that game audio is not at all similar to that of Hollywood movies. He states that video game audio is intended to be interactive, and must be engaging over countless hours of gaming, which seeks to provide an entirely different experience from a film soundtrack that is merely intended to engage the audience for 90-180 minutes.<sup>167</sup> For interactive music composers this is the fundamental difference between linear composition and an interactive gaming soundtrack, and what makes interactive orchestration such a challenge to create. The game must be programmed in such a way that it is able to anticipate, predict, and be aware of a player's actions in order to translate that into an auditory mood in real-time, but the soundtrack must still allow for a natural evolution in which music advances in a seemingly unscripted way. Ross refers to this as the audio adapting to environmental changes, giving a "portent of things to come", which is the essence of interactivity in game audio.<sup>168</sup> The idea is to build tension and suspense through the audio, and guide the emotion of the user seamlessly through transitions between game events. Kurt Harland says that these interactive considerations through audio can only be achieved by altering ways of thinking about musical structure and approaches to composition. He says this is integral in creating an emotional tone that follows the user's experience, making the environment feel more real and immersive rather than a linear, static composition. Instead of a composition flowing in a straight line from left to right, he suggests that interactive compositions should be thought of as a skeleton with several branching arms: "Only some of those arms are heard at any moment. The arms represent gamestates (character low on health, enemies nearby, etc.) and have to work well with each other if there's a chance of them playing at the same time. The branches have to shade the entire area of the song because you don't know when they'll be needed."<sup>169</sup>

There are some variations in how video game sound artists define interactive audio. In his article "Producing Interactive Audio" for *Game Developer Magazine*, author Mark Steven Miller suggests that the term "interactive audio" suffers from an identity crisis, and has come to mean less and less. He states that interactive audio implies more than linear playback; that the audio should feature reactive, responsive playback from audio drivers that are "aware" and can "respond" by shifting the music appropriately to the user's actions. Miller feels that audio can

---

<sup>166</sup> Belinkie, Matthew. "Video game music: not just kid stuff." *Video Game Music*. May, 2002.  
<http://www.vgmusic.com/vgpaper.shtml>.

<sup>167</sup> Ross, Rob. "Interactive Music...er, Audio." Gamasutra, 2001.  
[http://www.gamasutra.com/resource\\_guide/20010515/ross\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/ross_01.htm).

<sup>168</sup> Ibid.

<sup>169</sup> Harland, Kurt. "Composing for Interactive Music." Gamasutra, 2000.  
[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

only truly adapt if designers and composers undergo a major paradigm shift in how audio is conceived in order for audio tools to convey a sense of awareness of its own programmed responsive behaviours to the virtual world.<sup>170</sup> Published in the same October issue of Game Developer Magazine, Daniel Bernstein, in his article "Creating an Interactive Audio Environment", claims that sounds are affected by (and react to) environmental aspects of the scenes that are conveyed in the game, and this can only be properly achieved and executed by implementing a polished *Audio Object Vocabulary*, which is a method by which game objects communicate with each other and with the player.<sup>171</sup> Bernstein distinguishes between three separate types of sonic object interaction that make up interactive audio: direct, indirect, and environmental. *Direct* object interaction is a simplistic cause-effect relationship of sounds, for example, a character will cry out while dying from a gunshot wound. *Indirect* communication refers to the indirect method of object interaction in which something responds sonically as caused by something else that's happened in the game; for example, a character will breathe heavily when its energy is low. Environmental communication relays information about the soundscape to the player, atmospheric elements, which functions to reinforce a character's existence in a given environment that will sound automatically in a given context.<sup>172</sup> It is important to understand that interactivity is achieved by having the audio adapt to changes in the gaming environment, and implies more of an interaction between sounds and the actual state of the game rather than a direct interaction between the audio and the player.<sup>173</sup>

When surveying the literature written by video game audio designers it becomes apparent that the distinction between interactive and adaptive audio is increasingly unclear. Sometimes the terms are used interchangeably, while some will prefer one or the other to refer to audio that reacts to events initiated by the user in order for sounds to adequately adapt to gameplay. Both terms imply a soundtrack that readily adapts itself to the player's actions and changes in the environment and mood.<sup>174</sup> The Interactive Audio Special Interest Group (IA-SIG) describes an "Adaptive Audio Engine" in home computer game technology as an interactive tool that uses "special scripting methods to change the state of music dynamically throughout an interactive application" rather than the previous methods of "fading tracks in and out and jumping from

---

<sup>170</sup> Miller, Mark Steven. "Producing Interactive Audio." *Game Developer Magazine*, 1997.  
[http://www.gamasutra.com/features/sound\\_and\\_music/111497/interaudio1.htm](http://www.gamasutra.com/features/sound_and_music/111497/interaudio1.htm).

<sup>171</sup> Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in *Game Developer Magazine*, 1997.  
[http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm).

<sup>172</sup> *Ibid.*

<sup>173</sup> Harland, Kurt. "Composing for Interactive Music." *Gamasutra*, 2000.  
[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

<sup>174</sup> Pidkameny, Eric. "Levels of Sound." *VGMusic.com*, Video Game Music Archive: 2002.  
<http://www.vgmusic.com/information/vgpaper2.html>.

music segment A to music segment B.”<sup>175</sup> Rob Ross suggests that the term “Adaptive Audio” was a term used to describe a previously employed method of switching audio tracks. He goes on in his article "Interactive Music...er, Audio" to say that adaptive audio caused its own demise, in that the process involved the constant loading and unloading of several large audio files, thereby slowing the process of audio loading, and causing games to stutter. Ross suggests that interactive audio is a more accurate label than adaptive because “interactive audio is a technology designed to allow specifically created audio, placed in a given application, to react to user input and/or changes in the application environment.”<sup>176</sup> Although the distinction is still hazy in designer circles, the term *interactive* audio is used here to describe current audio technology in home console video games that allow sounds to adequately adapt to events in real-time during gameplay initiated by the user.

Regardless of which of the above definitions you wish to adopt to define interactive audio in video games, the process for applying interactivity in game audio is done through a series of *algorithms*, which is a set of computer procedures that are programmed to dictate specific sounds that the CPU will generate in relation to scenarios and gamestates. Algorithms are incorporated in video game audio to handle complex activities as a set of rules so that sonic events are automated accordingly. Keith Arem, game composer and sound designer for *PCB Productions*, states that game environments can incorporate hundreds of pieces of dialogue, music, ambience and sound effects, each with their own variable volume, pitch and positioning within the 3D audio environment.<sup>177</sup> He identifies the algorithmic process of audio design for video games as an obvious difference from film sound because interactivity requires that each sound be individually manipulated and programmed within codes to respond in real-time. Moreover, algorithms create an open-ended time flow, for which the designer cannot know in advance how long the player will remain in a given gamestate. This becomes one of the biggest challenges to audio designers in implementing interactive audio algorithms that generate music to be altered in real time rather than being programmed in advance to be triggered as fixed tracks or samples. Algorithmic composition is essential in interactive composition, in that designers input auditory cues and sound events into the game so that they will react sonically for the appropriate gamestate in real-time computing. Through algorithms, audio designers can arrange sound effects, soundscapes, and musical scores according to gamestates and scenarios so that the computer will perform the

---

<sup>175</sup> Alexander, Brandon. "Interactive Music: Merging Quality with Effectiveness" Gamasutra, March, 1998: [http://www.gamasutra.com/features/sound\\_and\\_music/19980327?interactive\\_music\\_quality\\_intro.htm](http://www.gamasutra.com/features/sound_and_music/19980327?interactive_music_quality_intro.htm).

<sup>176</sup> Ross, Rob. "Interactive Music...er, Audio." Gamasutra. May, 2001. [http://www.gamasutra.com/resource\\_guide/20010515/ross\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/ross_01.htm).

<sup>177</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.xvi

necessary tasks to make the audio adaptive and believable during the course of the game, thereby creating a greater sense of immersion for the player.

There are numerous computer music programs available to aid interactive composition. In The Algorithmic Composer, author David Cope comments on the wide range of computer composition programs that are made to generate seemingly “random” music using pattern-matching algorithms and assorted genetic algorithms.<sup>178</sup> There are also synthesis programs that allow composers to create and modify timbres and sonic environments (Loy, 1989), while sequencer programs allow the designer to control playback sequences (Puckette, 1991).<sup>179</sup> MAX is a prime example of a popular and widely available compositional program, which uses graphic interfaces, allowing composers to connect and program digitally represented synthesizer and sampler components.<sup>180</sup> Todd Winkler, in his book Composing Interactive Music: Techniques and Ideas using MAX, speaks about “interactive improvisational systems,” describing the art of composing interactive music in such programs as “music composition or improvisation where software interprets a live performance to affect music generated or modified by computers.”<sup>181</sup> Newer programs, such as *Reason*, *Rebirth*, *Recycle* and *Reaktor*, integrate sample triggering, loop production, MIDI sequencing and software synthesis into the compositional process.<sup>182</sup> Cope explains that interactive composition implies and necessitates an interdependent collaboration between composer and computer, and emphasizes that randomness and recognition depends as much on the listener or player and the context in which they are perceived interactively.<sup>183</sup> The composer or sound artist for interactive audio creates a piece of interactive music using algorithmic composition tools, and the audio designer then implements the score into interactive audio tools using parameters encoded to suit the shaping of the music in accordance with the state of the game.

Music is generally *event-driven*, which is a set of computer algorithms working in accordance with computer *artificial intelligence* (A.I.) to design responses to game events stimulated by the user. Event-driven music refers to audio scripting in which sound designers customize sound cues to correspond with audio events. For example, in a scenario that features a solitary enemy character entering a room, a designer would probably insert a cue for music of low intensity to play, while a room in which a whole army of enemy characters are waiting to ambush

---

<sup>178</sup> Cope, David. The Algorithmic Composer. A-R Editions, Inc. Madison, Wisconsin, 2000. p.14-15

<sup>179</sup> Ibid. p.15

<sup>180</sup> Ibid. p.246

<sup>181</sup> Winkler, Todd. Composing Interactive Music: Techniques and Ideas Using MAX. The MIT Press, Cambridge. 1998. p.4

<sup>182</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.214

<sup>183</sup> Cope, David. The Algorithmic Composer. A-R Editions, Inc. Madison, Wisconsin, 2000. p.38

the player would likely merit a cue for music of the highest level of intensity. Game A.I. is commonly used to describe how 'smart' the game is, and requires very detailed instructions and programming in order for it to function according to real-time parameters. A.I., signifying artificial intelligence, is a buzz word used by game designers and avid gamers to refer to systems within the game and gaming console that contain data corresponding to the perceived awareness and responsive logic that the game exhibits.<sup>184</sup> Games can and will incorporate several A.I. systems that in turn apply to different game elements (i.e. graphics A.I., sound A.I., etc). The term A.I. can have multiple connotations when speaking about video games, as game reviews often allude to A.I. as the perceived intelligence of virtual characters within the game.<sup>185</sup> Each component of a game will be encoded with specific A.I. data for every aspect of the game, which uses algorithmic functions to know what different components of the game are doing, and can react and respond using a set of programmed messages that have been cued by the designers to interact through gameplay at specific intervals. A.I. systems will communicate with one another within a game through an *Application Programming Interface* (API). The technical parameters increase in complexity the deeper one wishes to delve into the technology and its specific functions. Any further technological exploration into the hardware and software of game systems, with the exception of algorithmic processes, is outside the scope of this study.

An expert on musical composition using machines, author Robert Rowe outlines methods of algorithmic composition, combining the fields of music theory and cognition, computer-generated music, and artificial intelligence. Rowe, in his technical overview of computer music composition, *Machine Musicianship* (2001), deconstructs the theories and technological make-up of interactive musical algorithms. Microsoft's *DirectMusic* interactive audio program for use within the Windows *DirectX* package incorporates many of the algorithmic composition methods described by Rowe. The program can currently be downloaded for free, and makes interactive composition simplified and widespread for aspiring video game audio designers. Algorithms, Rowe explains, apply to different structures within the realm of sound when composing using machines, having multiple algorithms applying to such acoustic elements as pitch contours, time structures, pattern recognition, segmentation and scripting, multimedia installations, etc.<sup>186</sup> For example, algorithms directed specifically toward the processing of pitch materials address such components within 'symbolic pitch processors' as chord identification, amplitude computation, key induction, harmonic analysis, virtual pitch, etc. In order to create an algorithm that can

---

<sup>184</sup> Marr, John. "The Game AI Page." Game AI.com, 2003. <http://www.gameai.com/ai.html>.

<sup>185</sup> Reviews that speak about A.I. as pertaining to character traits within games include features in Gamespy, 2002: <http://www.gamespy.com/SOCOM/inside/gamefeatures.asp>, and BlueGamers, 2002: <http://www.bluegamers.com/ps2/reviews/mohf>.

control pitch contours that are context-sensitive, to interact in real-time with events created by the user, specific programming needs to be established and encoded into the programs by designers. Notes are stored and played back, combining musical elements that were composed and encoded by human composers and designers. Although algorithms for different games function differently according to the different parameters that are specific to the game, a frequency algorithm will typically incorporate a set of functions that will allow the algorithm to play the appropriate notes, chords and partitions corresponding to a game event. A *triad classifier* (three-note chord) is a code that is incorporated for the sake of chord recognition, which can alter the *tonal pitch class* (TPC) or *neural pitch class* (NPC) which, coupled with key induction, can be labeled and chosen by the algorithm that maps each pitch event into a harmonic point in relation to a game event.<sup>187</sup>

Some music tools employed by designers and composers include chord maps or "personalities" which can be used by the composition engine to generate segments, therein defining a type of road map for how the computer will place chords in a chord progression.<sup>188</sup> Separate algorithms function in conjunction with symbolic pitch processors, such as *neural networks* incorporating *scale processors* that will consider time structures such as *quantization* and *beat tracking* so that time intervals in the score will respond to tempo factors and dynamics of rhythm and duration.<sup>189</sup> Because more than one set of sounds can play at once in a game, tempo data is associated with short patterns called "motifs" that are short samples intended to be triggered by specific events.<sup>190</sup> Algorithmic composition must group musical events and recognition patterns together by means of *segmentation* and *pattern processing* to recognize intervals and segments within interactive musical structures.<sup>191</sup> Of course, memory capacity imposes necessary limitations that make the composition and implementation of a different audio track for every possible gamestate an utter impossibility.<sup>192</sup> Pattern processing and segmentation eliminates the need for a separate track for every gamestate, combining different segments at assorted stages to sound like entirely new sets of sounds. A basic parameter for pattern selection encodes different patterns and sequences into numbers from 1-100 called the "Groove Level," which imposes a number of variables to patterns; the more intense the gamestate becomes, the

---

<sup>186</sup> Rowe, Robert. *Machine Musicianship*. MIT Press, Cambridge: 2001.

<sup>187</sup> Ibid., p.9,44,46-47,61

<sup>188</sup> Hays, Tom. "DirectMusic for the Masses." Gamasutra, 1998.

[http://www.gamasutra.com/features/sound\\_and\\_music/19981106/directmusic\\_01.htm](http://www.gamasutra.com/features/sound_and_music/19981106/directmusic_01.htm)

<sup>189</sup> Rowe, Robert. *Machine Musicianship*. MIT Press, Cambridge: 2001, p.93,110-112,122.

<sup>190</sup> Hays, Tom. "DirectMusic for the Masses." Gamasutra, 1998.

[http://www.gamasutra.com/features/sound\\_and\\_music/19981106/directmusic\\_01.htm](http://www.gamasutra.com/features/sound_and_music/19981106/directmusic_01.htm)

<sup>191</sup> Rowe, Robert. *Machine Musicianship*. MIT Press, Cambridge: 2001, p.144-145,168,176.

<sup>192</sup> Harland, Kurt. "Composing for Interactive Music." Gamasutra, 2000.

[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

higher the Groove Level.<sup>193</sup> The auditioned segment will vary depending on what pattern has been assigned to a particular level, and designers can currently implement up to 32 possible variations for each pattern.<sup>194</sup> Rowe acknowledges the challenges inherent in making algorithms interactive so that A.I. will change its behaviour in response to the user's input. In order to achieve true interactivity through algorithmic composition, Rowe speaks of a "cross-fertilization between music cognition and machine musicianship," which requires the "intersection of music theory, cognitive science, and artificial intelligence."<sup>195</sup> He also acknowledges that designers must be aware and make use of *schemata*, which is defined by Albert Bregman as "a control structure in the human brain that is sensitive to some frequently occurring pattern, either in the environment, in ourselves, or in how the two interact."<sup>196</sup> Schemata are an important consideration when designing algorithms for interactive audio because they are a way of anticipating organized responses to situations. Algorithmic programs must be modeled to incorporate music cognition to know when certain patterns of activity will be encountered, requiring certain audio sequences (that are stored in memory) to be activated.

Interactive music, or "*i-muse*" as it is sometimes referred to in the video game industry,<sup>197</sup> must incorporate notions of "emotional transduction" that are designed to evoke a feeling of dialogue for the player through gestural changes in the interactive musical score.<sup>198</sup> One of the recent successes of interactive music is the incorporation of *algorithmic synthesis* methods that can add dynamic variation to a *wavetable* sound. Denis Labrecque of *Analog Devices Inc.* explains "the combination of wavetable, physical modeling, DSP and any of the other algorithmic synthesis techniques within a single sound environment can lead to a flexible, interactive, larger-than-life listening experience."<sup>199</sup> The algorithms provide a sense of variability that helps to make sounds unpredictable, and thus more believable in the virtual context. Audio programmers fashion audio algorithms and sound software authoring tools, giving them the ability to interpolate between sequence states, track states, and voice states to provide audio designers with

---

<sup>193</sup> Hays, Tom. "DirectMusic for the Masses." Gamasutra, 1998.

[http://www.gamasutra.com/features/sound\\_and\\_music/19981106/directmusic\\_01.htm](http://www.gamasutra.com/features/sound_and_music/19981106/directmusic_01.htm).

<sup>194</sup> Yackley, David. "Microsoft DirectMusic: Creating New Musical Possibilities." Microsoft Corporation, 1999.

[http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnmusic/html/dm\\_nmp.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnmusic/html/dm_nmp.asp).

<sup>195</sup> Rowe, Robert. *Machine Musicianship*. MIT Press, Cambridge: 2001, p.235

<sup>196</sup> Bregman, Albert. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MIT Press, 1990. p.401

<sup>197</sup> Belinkie, Matthew. "Video game music: not just kid stuff." *Video Game Music*. May, 2002.

<http://www.vgmusic.com/vgpaper.shtml>.

<sup>198</sup> Lewis, George E. "Singing the Interactivity Blues." *Western Front Magazine*, 1999.

<http://audiolabo.free.fr/revue1999/content/lewis2.htm>

<sup>199</sup> Labrecque, Denis. "Altering Algorithms to Create '3D' Sound." *EE Times*, 2001.

<http://www.etimes.com/story/0EG20011218S0047>.

interactive control of the game music.<sup>200</sup> Audio scripts are compiled and played back using "control logic" commands that are algorithmically constructed to tell the game how to react to user activities and which music to play depending on game events. Different sets of sound samples are incorporated much like "sound fonts" that can be loaded for different scenarios, and can be triggered at different intervals, keeping the audio fresh throughout the game.<sup>201</sup> By design, interactive music for video games will conform to the mood of the player and the game setting. The music is designed to flow seamlessly and continuously throughout the game's progression, with computer A.I. constantly adding or subtracting layers to give varying levels of intensity for the same piece of music. Truly interactive music will be prepared for any possibility, swinging with each turn of events to provide a feeling of an adaptive soundtrack that evolves as the player advances the storyline. Robert Rowe acknowledges that this interactive improvisation is perhaps the greatest challenge to implement, as the "machine must contribute a convincing musical voice in a completely unstructured and unpredictable environment."<sup>202</sup> The difficulty is in making the program able to rely on its own music cognition or programmed common sense to "derive structure from what it hears, and impose structure on what it produces in response."<sup>203</sup> When all of the elements work properly in conjunction with one another to create a dialogue with the player, the outcome is a complex, random, ever-changing and seamless plethora of sound within the artificial environment that propels the game forward and enhances realism within the artificial context. Aaron Marks passionately describes the result of an effective interactive music sequence in The Complete Guide to Game Audio.

Music will be slow and surreal as a player is exploring a new environment. If the game character moves from a walk to a run, the music will also keep pace by increasing in tempo. As danger approaches, the music will shift to increase tension, then to an all out feeling of dread as the bad guy appears out of the shadows. The player chooses a weapon and begins the attack as the music shifts again to a battle-esque theme to fire the player up. As the player gets hacked to pieces by the bad guy, the music will turn dark conveying that the end is near. But, as the hero gets his second wind and begins a determined counterattack, the music will morph into a triumphant flurry as he imposes death and destruction on the villain's head...the feeling is indescribable.<sup>204</sup>

Interactive game audio involves new ways of thinking about sound, music, design and composition. Levels of interactivity inherent in a game's audio structure will depend on the genre of the game itself, in addition to the storage capacity, RAM allocation, disc speed and other

---

<sup>200</sup> Patterson, Scott. "Interactive Music Sequencer Design." Gamasutra: 2001.  
[http://www.gamasutra.com/resource\\_guide/20010515/patterson\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/patterson_01.htm).

<sup>201</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.277

<sup>202</sup> Rowe, Robert. Machine Musicianship. MIT Press, Cambridge: 2001, p.277

<sup>203</sup> Ibid., p.277

<sup>204</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.191-192



technical considerations and limitations in specific home console game systems. Although interactive techniques are far from formulaic, and vary amongst different designers, developers, publishers, programmers and game platforms, the strategies for making game audio live up to its interactive potential and user expectations are standard goals in the development process in expanding the overall effectiveness of the game. The potential for total immersion, viable in large part because of audio interactivity, could possibly make the experience of playing interactive entertainment fully engaging and encompassing, surpassing that of watching television and movies. The participatory experience of modern home console gaming, coupled with the meanings and sensations evoked by interactive audio, provides promise in allowing for full captivation in an interactive entertainment experience. New advances in audio technology have brought sound in video games to the forefront of the experience, improving the overall quality of the aural event. Interactive considerations in audio are intended to reach the player on a subconscious and emotional level, often signifying subtle cues that are felt rather than shown, embodied through aural perceptions in ways that visuals cannot hope to induce. In the next section I will analyze the sonic elements that are integrated to make up particular game audio environments, and will be applying interactive considerations to evaluate the potential for interactivity in the audio design of those specific games.

*Clearly, music, sound, and images responsive to user input will increase the user's level of interest and the feeling of participation.*<sup>205</sup>

-Todd Winkler

*By design, interactive game music will adapt to the mood of the player and game setting... If you were to watch a recording of another person playing the game, a good interactive soundtrack would seem as if it was specifically scored to the scene, similar to a linear movie... Game players are not predictable. We don't know when they will walk, run, hide, enter a new room, meet the bad guy, draw their weapon, or do any of the hundred possible actions that can happen during a game. But, truly adaptive music is prepared for any possibility. Ensuring the music can transition naturally is what makes it work.*<sup>206</sup>

-Aaron Marks

Video Game Audio Designer

#### **4. Principles of Design: Hearing Is Believing**

Having surveyed the literature about the history and theories of video games, interactive audio technology, algorithmic composition, interactivity, having provided a detailed overview of acoustic terminology, psychoacoustic and soundscape theory in the first three chapters, this section aims to provide some insight into principles of sound design employed by the professionals at Electronic Arts (EA) Inc. In order to support my research documented in the previous chapters, and to provide a design insight for the analyses to follow in the case study chapter, the information gathered from designers in the world's leading independent developer of interactive entertainment software will serve to provide an inside perspective about interactive audio considerations in modern games employed by the professionals in the industry. The statements made by the designers will serve as information distinctive from literary research to further substantiate the research conducted, as well as to support the findings and hypotheses to be formed in conducting the study in the next chapter, contributing to a well-rounded discussion of audio in interactive entertainment. The names of the designers have been fabricated to respect their confidentiality and to ensure their anonymity. The questions posed and information obtained in this section have been designed to uphold the conditions stipulated in the Non-Disclosure Agreement (NDA) upheld by EA. The observations and information contained in this chapter are based on my interpretations of the information conveyed by the designers in the six, independent interviews conducted, and my aim is to incorporate their statements and opinions in an objective

---

<sup>205</sup> Winkler, Todd. Composing Interactive Music: Techniques and Ideas Using MAX. The MIT Press, Cambridge. 1998. p.297

<sup>206</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.191-192

manner in order to offer educated perspectives of interactive audio and its function from those who design and implement techniques of audio interactivity. Information about the specific tools and techniques used by designers at EA was not divulged in compliance with the NDA, and any technical information about interactive audio systems was obtained and documented prior to the interviews.

The professionals involved in interactive audio composition and implementation in video games generally prefer to adopt the title of sound artist or sound designer, rather than that of their specialization within the audio production team. In the art of creating sound for motion pictures, the specializations of the workers in specific positions of the sound team are clearly defined, and often segregated. Because sound artists for video games generally occupy a wide range of specializations, requiring a great deal of audio multitasking, the term “sound artist” or “sound designer” generally encompasses a number of activities, such as sound recordist, audio programmer, Foley artist and sound editor.<sup>207</sup> A team of audio designers often work together from the earliest discussion stages at the onset of a game’s production, and work closely together into the final “beta” stages of production to ensure that the components of interactive speech, music and SFX work together seamlessly within the environmental construct, to allow for the player to encounter an engaging and encompassing aural experience. Many come from backgrounds of computer science, engineering, and commonly backgrounds featuring musical interests and a passion for video games. As musicians, many of the sound artists are familiar with techniques of composing, arranging, copying, editing, engineering and producing.<sup>208</sup> The instruments employed by video game artists consist of sound libraries, recorded instruments, licensed songs and samples, and any other acoustic and electroacoustic sources available to them. Of course, designing interactive audio for real-time playback requires a different approach from conventional music composition and arrangement, as articulated in the previous chapter of this thesis. Audio hardware and software continues to advance, and audio tools and algorithms are progressively reinvented to evolve with technological advancements. The specific tools and techniques used by the designers at Electronic Arts Inc. will not be divulged here, other than to state the obvious fact that sound designers at EA are incorporating the newest and most advanced interactive audio technologies available, as can be heard in games like *SSX Tricky* and *MVP Baseball 2003*.

The sound artists were each chosen for their role in the audio design process, each playing a key role in the interactive audio design of current game titles. *GD* works as an ‘audio

---

<sup>207</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.222

<sup>208</sup> Ibid. p.222

lead' at EA, whose job entails deciding how to go about meeting the audio needs of games to comply with the demands of producers, assembling an audio team to adequately disperse specific tasks, and to adapt the appropriate audio tools and software to suit the needs of the audio team. His job is to determine what needs to be evaluated for specific game audio production, to program audio software, and to design AI parameters to obtain the desired effects from the audio systems. GD has a background in interactive MIDI instruments, and a degree from SFU. *JMR*, also an audio lead, works with producers, programmers and designers to ensure that the goals of the audio production are being met. He must ensure that the audio design elements come together properly to suit the needs of the game, and works with an audio team, specializing in music. *JMR* has a musical background, having played instruments for numerous bands and having worked with both analog and digital music tools. *AZ* works as a sound artist at EA, specializing in, but not limited to, sound effects production. *AZ* studied electroacoustic music and sound aesthetics, and has worked on a wide array of audio projects involving music composition, sonic environments and radio documentaries. *JMAC*, also a sound artist, works on the speech development team, designing speech commentary for sports titles. *JMAC* holds a degree in computer engineering from Queens University, and developed a keen interest in digital audio leading up to his position at EA. *IM* is a lead programmer on the audio tools group for EA. To his credit, *IM* has designed a number of speech tools for sound artists at EA, creating audio software to allow artists to adjust and control parameters to their specifications. *IM* was originally hired by EA as an "electronic sound synthesist," and began by programming the previous audio technology of sound chips into consoles. *IM*'s background includes a degree from SFU, as well as having worked in various projects involving studio work, sound composition, programming, synthesis, and real-time interactive music programs. Last, but certainly not least, *JMO* is a designer and in-house composer of interactive music at EA. *JMO* has composed and crafted interactive music for a number of successful game titles, including *SSX Tricky*. His background is entirely musical, having toured with Ozzy Osbourne and Marilyn Manson as a keyboard musician with several successful bands such as the popular metal band Fear Factory. *JMO* composes electronic music for EA, primarily with synthesizers, and is one of only five music composers at EA to work exclusively with music. The six interview subjects are evenly dispersed in their specializations, and were chosen to fit within the analytical framework of this study: *IM* and *JMAC* specialize in speech components, *JMO* and *JMR* in music, and both *GD* and *AZ* work primarily with sound effects. Each of the artists was interviewed separately to preserve unbiased responses to a series of questions pertaining to interactive audio. Their key responses are summarized and documented as the basis for this chapter, for the purposes of understanding interactive audio considerations

from the point of view of the artists who constructed and shaped the interactive sonic environment.

#### 4.1 Interacting with Interactive Sound Artists

The questions were posed in a conversational manner to create a more relaxed atmosphere in order to facilitate an open-ended discussion, and to obtain the most information possible from the interview subjects within each hour-long dialogue. The responses were generally varied, and I have grouped them in an orderly fashion to serve as informative contributions to the interactive audio discussion, maintaining the analytical framework to be modelled in the next chapter.

To instigate a thought-provoking discussion, I posed the question “*How do you go about making game audio that is interactive, realistic, exciting and immersive, keeping the player coming back for more?*” The question was designed to bring about a discussion of interactive considerations in audio design for the purposes of promoting interactivity, and more specifically how this interactivity is achieved through audio. As IM suggests, the level of interactivity will vary, depending on the game genre; that audio in some games is geared more toward improving spatial characteristics to create a vivid soundscape to immerse the player than making sounds adapt to the player. IM states “Making it different every time, and making [sounds] appropriate is important,” which requires designers to be creative in applying the appropriate input parameters, or “patches” to make sonic components adaptable depending on the given scenario. For example, IM gives an example of how interactivity might be emphasized in a sports sim basketball game, by having the loudness, pitch, and rhythm of the sound of the ball bouncing on the court change in relation to the height of the character, the character’s ball-handling skills and player stamina. Each sound needs to be scaled and adjusted ahead of time by designers so that sample outputs will apply appropriate sounds and their parameters to adequately suit the gamestate. Input parameters will also be programmed to adjust for the physical environment of a given venue or soundscape, so that sounds will react to the environment in which they occur. In the above basketball example, the sound artists will decide what input parameters need to be applied so that “the mapping and tweaking of those variables can control the outcome based on those inputs.” Designers need to account for any missing parameters to make the sounds believable within the context in which they occur, such as spatial parameters within the environment. IM emphasizes that it is all a method of “parameterizing” all of the sounds to correspond with gamestates, events, and the functional, sonic environment so that various input controls can be altered in real-time by game AI and audio tools in relation to a player’s actions.

In response to the above question, JMAC maintains that the game audio adapts by responding with appropriate sound samples that are encoded by designers ahead of time to respond in real-time. In the case of speech for sims, the “input parameters are passed through the speech AI to have the appropriate speech samples chosen to play for the proper gamestate.” What is important is that the speech provides feedback for the player, to signal whether the player is doing well or not, and to provide information (through subtle hints or direct cues) on how the player can improve gameplay. According to JMAC, the goal of interactive speech is to create a realistic dialogue that flows with gameplay; to provide the experience and make the player believe that what is being spoken about applies to what the user is doing at any given time. Furthermore, JMAC asserts that intelligent speech to correspond with a given action in real-time is what is strived for, more than just speech that describes an event and is triggered after an action is performed; “speech carries you through the action with ongoing, relevant and intelligent commentary.”

Musically, JMR claims that the approach is to vary the musical experience according to the game genre, through “momentum-based criteria”, so that the score evolves as an “emotional building process.” JMR distinguishes between game genres, upholding that some games follow conventional song structure, while others need interactive rises and falls within the score. Interactive or not, JMR maintains that the music must follow the experience and support the mood of the game. For example, in sports sim games the music isn’t generally a foreground element, and is meant to flow as it would in a real, live sporting event. Music designers study real sports telecasts as a prototype for what they need to implement, in order to replicate that experience. As the motto for EA Sports states: “If it’s in the game, it’s in the game.”<sup>209</sup> JMO comments on his own experiences in crafting interactive music for fantasy sports games, saying that “the music remains more or less constant throughout [gameplay], but shifts during game events” to symbolize or incite a modification in play style. “Each game must be examined ahead of time to determine what the game modes are, and how music can function interactively to enhance and guide [the user] through gameplay” JMO maintains. One of the ways music can work as an interactive element to augment the atmosphere of the game during an event is to increase the tempo, making the music more energetic to create a sense of urgency, thereby heightening the feel of the game. “This is like the musical impact that happens in a Hollywood movie,” JMO continues, shifting the focus to role-playing games “where the music follows the mood of the experience and adapts to transitions in gamestates...it can create dramatic tension and represent uncertainty.” This is the way game music should function, according to JMO, that it

should be able to tell a story where appropriate, to “provide an ever-shifting mood for the player in different gamestates.”

In terms of SFX, AZ articulates that interactive audio has to be both random and logical, but must also defy the user’s expectations in order for it to be truly effective. “Intelligence,” AZ states “is reflected in the element of surprise, [achieving a] balance of unexpected outcomes and cues for the player to take specific actions, and to indicate that they’re on the right track.” The job of sound designers is to modify the sonic environment so that it “resonates with the player.” If effective, the player recalls consequences of specific actions taken in the gaming environment, but must be ready for unexpected shifts as gamestates change. GD further elaborates on this, explaining that sounds must be layered and be encoded with the appropriate input parameters so that the audio AI can anticipate every possibility, and recall the appropriate sounds for real-time playback. “Sound effects,” GD proposes, “[must be] organized as tasks so that sounds will respond to the appropriate gamestate”, thereby building an “event system” through designing “module banks” that incorporate real-time components. This is done in various ways depending on the game, according to GD, either having sounds play back interactively as “player driven” components, or in response to gamestates or environmental cues with “physics driven” sounds. GD explains that electronic messages will be passed through SFX engines to layer sounds from specific module banks to play back the appropriate sounds with suitable modifications in relation to a player’s actions, such as compression, filtering and effects patches to vary sounds according to what a player does. Physics-driven sounds still function interactively, he maintains, although this is done through techniques of “surface mapping”, which involves “tagging” surfaces of the surrogate environment with a parameter number that real-time audio algorithms will recognize, thereby playing the appropriate sound sample and applying corresponding modifications in real-time, based on physical parameters in the game environment. Audio designer for the game *Halo*, Marty O’Donnell, explains the process of tags and their potential to instigate random sounds.

Every piece of raw audio data is called a “soundfile” and the set of instructions that organizes and determines how the soundfile is to be played is called a “soundtag”... The most important feature of a soundtag is that it contains enough permutations and the proper randomization so that players don’t feel like they’re hearing the same thing repeated over and over. Even the greatest and most satisfying sound, dialogue or music will be diminished with too much repetition. It’s also important to have the ability to randomize the interval of any repetition.<sup>210</sup>

---

<sup>209</sup> ©2003 Electronic Arts, Inc. *MVP Baseball 2003*. (video game)

<sup>210</sup> O’Donnell, Marty. “Producing Audio for Halo.” Gamasutra.com, 2002.  
[http://www.gamasutra.com/resource\\_guide/20020520/odonnell\\_01](http://www.gamasutra.com/resource_guide/20020520/odonnell_01).

This result is coupled with spatial parameters so that sounds play back in relation to a player's position within the environment, giving the player a sense of location within the artificial soundscape, and augmenting the realism through diverse sounds. Designer Will Davis, whose video game audio credits include music and sound effects in such titles as *Driver*, *Nightmare Creatures 2* and *Battlemorph*, offers his simple approach to making interactive SFX work in video games: "The goal is simply to create audio...that has the desired effect on the player...it has to 'feel' right...It doesn't matter if the sound is realistic or over the top or even completely wrong. If it works in the game and sounds and feels right, then it works."<sup>211</sup>

I proceeded to ask how "randomness" in audio is achieved, and about methods to make audio seem random to the player, crafting sounds that are unpredictable, and therefore more believable. IM responded by stating that layering is a main strategy to making sounds seem randomized. Speech commentary in sim games, for example, must sound like the commentators are reacting appropriately to the events. Speech must, therefore, sound fluid and natural, and believability and intelligence is perceived when speech samples are appropriate and non-repetitive. IM specifies that speech in sims follow the game action rather than directly affecting the pacing of the game. He maintains that "speech samples cannot say everything, nor should they." The importance is on modelling real conversations by real sports commentators, to be scripted and arranged so that the appropriate speech samples interject at the right moments, commenting on the game in real-time. Painstaking effort is taken by speech designers to ensure that important game events are spoken about. The only real concern, as IM puts it, is when a significant game event transpires and an event is not acknowledged, which can make the commentators seem ignorant, and calls attention to the artificial nature of the game environment. There can be different levels of interactive audio within the game, according to IM, who gives the example of NIS "mini-movies", that are by nature non-interactive, but can feature adaptive, non-scripted sounds, such as spatially-specific crowd sounds, that will react with panning parameters according to where the virtual camera moves. He cites an example of this in an intro sequence of *SSX Tricky*, where the camera flies in overhead to focus on the starting gate from behind the first player's character. The player has no control over such sequences, however the crowd sounds that encompass the player and shift with the camera's movement are unscripted, and correspond to tagged mapping parameters in an interactive way, relating to the player's position within the environment.

---

<sup>211</sup> Marks, Aaron. "An Interview with Will Davis." Gamasutra.com, 2000.  
[http://www.gamasutra.com/features/20000901/Davis\\_01.htm](http://www.gamasutra.com/features/20000901/Davis_01.htm).



JMAC states that interactive speech is primarily driven by the player's actions rather than by environmental aspects. He further articulates that random changes are made possible by advances in console memory. The game memory keeps track of which samples it has played by having parameters encoded into the speech AI so that the game will not play a given sample again until other samples have been played in between. JMAC calls these "intervening samples" a way of cycling groups of samples over a number of consecutive games to avoid repetition of speech commentary. The PS2, for example, will save its samples in memory until the console is turned off. He admits that the interactive audio experience will be different for a player who turns off the console between games than for the player who saves games in progression, or plays consecutive games back to back without deactivating the console. Players who turn off the console between games "take the power out of the hands of audio designers," in that any variations that have been programmed into the interactive speech AI engines are wiped clean, and the tools must start from scratch. Even in this case, JMAC says that designers do make efforts to maintain the random element by having the speech AI randomly select speech comments in real-time from a group of appropriate samples for a given action. In all probability, the speech will be different, but acknowledges that this "random seed" approach does not guarantee variety, which can detract from the realism of the interactive dialogue created by the speech AI.

In terms of randomness in music, JMR explains that this is done through various music streams that are layered in real-time by music algorithms. Streaming is essentially playing a track from a CD, but is combined with randomized sound events that are stored in RAM, which allows designers to dynamically change the layers that will overlap at specific times over the song. The element of unpredictability is thus created through a process of layering music tracks that are synchronized in music tools, incorporating beat-driven sequences to ensure that layers all conform to the same tempo in compliance with the pacing of the game, as in *SSX Tricky*. JMAC distinguishes between this and the techniques of scoring orchestral music for games like *SOCOM*. Orchestral music is more difficult to make interactive, according to JMAC, because it must follow more ambient-driven transitions. Music functions more as an event-based soundtrack in this case, which entails programming AI and environmental input parameters that influence musical changes ahead of time through software treatments instead of in real-time. Parameters such as frequency, volume, panning and filtering are all implemented in the early design phase, and are programmed to respond to specific game events rather than adapting to a player's actions in real-time. While this may be limiting, JMAC acknowledges that it can be effective, and can function interactively as music that shifts along with gamestates for an immersive and engaging experience.

JMO insists that music has to be written interactively, and be sculpted for the interactive experience rather than borrowing pre-arranged music tracks to be deconstructed and integrated for the purposes of game scoring. JMO goes on to discuss how, in *SSX Tricky*, players know something significant has happened in the game by *feeling* subtle shifts in the score. For example, “as the tempo downshifts and instruments fall out of the mix, it signals to the player that something is not being done properly, or to its potential.” On the other hand, “when the tempo increases and instrumental layers begin to build, it signals that the player’s style is improving.” This is the idea behind interactive music designing techniques, in that players may not pay active attention to how the music is changing, but they feel something is happening through the music emotionally, without necessarily knowing that different layers are occurring in real-time. Random variables have to be programmed into the AI to create a random impression for the users that the soundtrack is ever-changing in relation to their evolving style, creating a personalized, individual, interactive experience. When I asked how this is made possible, he explained that *SSX Tricky* had three random factors that could affect the music during gameplay: player style, ranking and tricky boost meter. JMO explains how the soundtrack builds itself in real-time:

As the player gets better in the game, scoring high in points, riding in 1<sup>st</sup> place, or reaching maximum boost on the tricky meter, the music gets more intense. The audio AI takes [into account] musical data, using regions, and other input parameters to play back the songs to reflect how well the player is doing in the game. The same race could be played dozens, or even hundreds of times without the soundtrack doing exactly the same thing... The song will layer differently in real-time to give a different mix of the song each time. If the player gets to be an expert, the “top level” mix will play the majority of the time. To avoid repeating this, I had random patches built in to change the musical sequence in case the player goes down the same course in the same way, to mix it up a bit each time. A good interactive music system is important to ensure that the same score doesn’t keep repeating. It has to seem different...

Designer for *Anarchy Online*, Bjorn Arve Lagim, explains how interactive music creates an ever-changing score in his article for Gamasutra.

Each sample has multiple transitions and will rarely move into the same next sample twice, and so the music is never played the same way twice. In linear time, *Anarchy Online* contains about two hours of music, but when played back with the [music] tool, the amount of unique music is almost infinite.<sup>212</sup>

AZ acknowledges that randomized events involving sound effects are difficult to achieve, but are done in such a way that “DSP reacts on the fly by combining acoustic functions to vary each sound”, not just to conform to triggered sample playback. One method of incorporating this

---

<sup>212</sup> Lagim, Bjorn Arve. “The Music of Anarchy Online: Creating Music for MMOGs.” Gamasutra.com, 2002. [http://www.gamasutra.com/resource\\_guide/20020916/lagim\\_01](http://www.gamasutra.com/resource_guide/20020916/lagim_01).

would be in the form of a sports sim crowd sound, which he describes in detail as documented below.

Crowd sounds resemble a granular synthesis of crowd-stream loops. The crowd streams overlap, but they are also ever-changing by creating timbral shifts in real-time through overtones that vary the random sonic picture. [...] What the player does, the nature of the gamestate, and the context of the game event will affect the audio mix, taking on different characteristics depending on the input variables. For example, the crowd's reaction will morph and change the... auditory gamestates to adapt to a shift in the game event. The transition will [blend] seamlessly, as the auditory gamestates are sound objects that change according to parameters that are, for the most-part, predetermined ahead of time, which morph through sonic events.

The idea, as AZ puts it, is that “psychoacoustically, the player must perceive change,” and sound designers use signal processing parameters to create this ever-changing perception, in order to alter the sound each time it is heard in a different context. As long as the sounds work within the sonic environment to immerse the player in the game, then the random variables are a success.

GD was adamant that “sound effects must feel and sound physically correct” within the context of the game. He gives an example of *Sledstorm*, a snowmobile racing game, in which engine sounds are realistically created to replicate how they would sound in the virtual sonic environment. The engine sounds are contained in a module that adds layers based on how the engine RPM in a “two-stroke snowmobile in a fantasy world, made up of junkyard parts, going 300 miles an hour” would sound. GD explains that parameters such as “acceleration, velocity, in-air compression, engine revs and landing kickbacks” are all accounted for in the auditory recreations of extreme snowmobiles in the game, and sonically, they must physically respond to the fantasy environment. As GD clarifies, “it has to physically respond. It has to wiggle and snarl, and do whatever a particular device would do in that fantasy world. It has to have a physical grounding for the player to buy into it; it has to physically make sense.” GD elaborates on how the complex sound effect of crowd ambience is implemented in order to create seamless, random shifts in order to maintain a sense of unpredictability for users.

A crowd sound might use about ten simultaneous voices, incorporating loops, samples... offset panning and stereo streams along with random actions that will play now and then, based on an internal timer within the environment. An example is when you'll hear a car pass every so often, or a foghorn blast in the background... They're all gradual transitions that are seamless and seemingly endless. [...] There are also RAM loops that create a curtain of more specific crowd responses to overlap the ambient crowd mumbling; then there are event sounds that occur based on direct player interaction. Like when the player makes an exciting move, it will trigger a huge cheer that adds to the ambient loops. The cheer will have other voices blended in to create this realistic crowd atmosphere... The loops will continue playing and modulating

in real-time from the time the game begins to the point where it ends or is turned off.

GD's vivid description is an ever-modulating and constantly overlapping mass of sporadically cycling and periodically variable loops that flow together seamlessly to anticipate the player's every move. It almost functions as a living, breathing entity that ebbs and flows with player actions. GD concludes his discussion by adding that players cannot notice the transitions or loop points because loops never overlap the same way twice. Sounds within the crowd are also treated with parameter patches, tables, and "randomized event scalers" (that have variants depending on a character's energy, for example) that will scale the volume and pitch (etc.) of a sound accordingly. The SFX AI system must be tuned and programmed with "logic" data to "take whatever the game is spitting out." The audio designer must evaluate every scenario or possible gamestate, and must program the AI to respond by triggering appropriate sounds according to Groove Levels and gamestates.

Next, I asked about audio cues that are implemented to evoke a response and signal the player to take specific actions in the game. For speech, JMAC stresses that speech cues can either be obvious or subtle, depending on the gamestate. In tutorial game modes (that demonstrate techniques for the player to practice before attempting gameplay) speech samples are designed to speak to the player, and directly instruct them. These sequences are often non-interactive, and feature scripted speech, as in Harold Reynold's tips on improving gameplay in *MVP Baseball 2003*. "Speech cues are subtle for the most part in speech commentary for sims," explains JMAC, emphasizing the effectiveness of suggestive speech to coax the player to adopt a different style of play. Without talking to the player directly, comments in speech commentary will reflect a player's actions within the match itself. This is often done with banter between the two commentators to signal something indirectly to players about how to improve their play style. For example, in *MVP Baseball 2003*, the colour commentator might say "There wasn't much velocity on that fastball. We're used to seeing this pitcher have far better command of his pitches." The implication that the character's performance is lacking functions as an indirect signal that the player should make adjustments to the style of play. A paralinguistic inflection, such as commentators getting excited about a play, will function as a reward, glorifying player actions without telling players they did well.

IM continues on this premise, saying that different layers of speech are often incorporated to hint at something the player did as a subtle cue. In *SSX Tricky*, for example, the character of Eddie will shout "Woohoo!" to signal his satisfaction at having accomplished a trick. IM explains that this is scripting, but it is consciously done in order to relate speech comments to the level of

the player, and can be designed both to provide direct feedback and enhance the fun. Speech scripting requires that both of these functions are met. He goes on to say that subtle tips and hints in speech for graphic adventure games are a prominent part of speech design, because players can often get stuck and need guidance, but they don't want to be told what to do.

JMR suggests that musical shifts can also provide subtle cues about how players might adapt their style of play. He explains that in FPS games musical shifts provide a double function in that musical build-ups not only serve as cues, but also provide tension and atmosphere. Aaron Marks discusses interactive considerations in music that act to set up upcoming confrontations and cue subtle warnings to players in FPS games. "The visuals don't typically change until the bad guy actually shows himself" Marks states, "the audio becomes all important in this instance. Without it, the player doesn't have any foreshadowing and won't know to take out their weapon or get out of the room until it's too late."<sup>213</sup> Interactive music cues, in this case, play to the player's intelligence, and allow them to incorporate their auditory senses to adapt to shifts in the game. Music, according to JMR, has powerful interactive potential to give subtle direction that plays on feelings rather than words. "True interactivity," he states "plays on emotions rather than logistics...Musical arrangements can add or subtract instrumental layers to signify something to the player through dramatic, musical shifts." JMR says that improvements are constantly being made to shift the musical arrangement in real-time, altering the bar structure in subtle, sophisticated ways. "Musically, we can alter section on the fly, and freshen up a piece by changing it in real-time through random access memory to create a dynamic shift." JMR stresses that the type of musical cue will depend on the situation, which takes much planning and evaluation. In fantasy sports games, a rhythmic overlay can signal the player to punctuate a trick or move, or the melody can change dramatically to signal to the player that something significant has occurred in the game, augmenting the gamestate. He explains that this technique should not be overdone.

There should only be a couple of unique [sonic] events that are hard to achieve in the game that should be dramatically emphasized in the music because it takes the game experience to the next level. There are generally so many things going on in the audio that there has to be a key moment where the game event stands out over the rest of the audio environment, so that it doesn't just become noise that gets lost in the overall mix.

JMR remarks that this takes much planning in the early stages of the design phase to make sure that achievements by the player will be superimposed, and have a unique musical treatment.

---

<sup>213</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.190

JMO acknowledges that there are times in the game where designers may want players to notice transitions, for which they implement triggers or add digital effects to the music to make the score more interesting. This can be done with jarring musical soundmarks, like the “It’s Tricky” theme that is modulated with effects once a player reaches a heightened gamestate, or can be done in a more subtle manner. JMO explains that one possible way would be to layer two separate beats that are beat-synched within the same tempo to create a more vibrant, percussive sound. Another example would be if a player goes through a tunnel, the music could reflect this with a “tunnel variation” of the original theme, which would be reverb-treated, thus creating a subtle shift to suit a transition in the game environment. JMO warns that dramatic shifts in the music should be avoided when there is no corresponding game event. “If the music changes when nothing happens, it will be lost on the player.” This can obscure the meaning of the musical transition, and could result in dramatic shifts during game events, losing their significance. “It has to make sense,” according to JMO, “you don’t want music that has nothing to offer.”

GD cites an example of interactive music cues being implemented through dynamic shifts in the fantasy sports wrestling game entitled *Def Jam*. He explains that the introductory theme music incorporates a heavier bass and rhythmic overlay during a fight event, which further intensifies in the game’s “blazin’ mode”. In this heightened gamestate, the music switches to a fast-paced, high-pitched variation of the song, coupled with a “swish” sound effect that highlights key moves during the event. The music crossfades back into the regular theme once the fight event subsides, creating a smooth transition. Dynamic effects are added in real-time, such as record scratch SFX and high-impact sounds of glass shattering, randomly inserting themselves to create larger gestures for sonic events. GD gives another example in the same game of a more subtle shift that occurs when a player leaves the ring after a fight. The music will continue, but will be filtered to sound muffled and distant, creating the impression that the music was left behind in the arena through considerations in the acoustic design of the virtual environment. This gives indications about the surrogate environment that envelopes the user, functioning interactively to follow the player’s actions and experiences.

One of my most pressing questions was the issue of real-time, reactive audio. I asked each of the designers a series of questions about audio interactivity, enquiring “*Does the audio react and respond to gameplay?*”, “*Are the audio drivers ‘aware’ of what’s happening in the game?*” and “*Can the user affect the audio performance of the game in real-time?*” The responses were substantial, and gave me insight as to the potential for audio interactivity in different game genres. The ‘awareness’ of the game audio is in the form of speech, music and SFX AI that have

variable parameters, randomized events and corresponding samples, and encoded “prioritizing” information for it to respond adequately to every possible scenario.

In terms of speech, JMAC speaks of the importance of encoding parameters into speech AI so that it ‘knows’ not to repeat samples unnecessarily. Console memory is now becoming sophisticated to a point where game sound is saved in memory even after the machine has been turned off (Xbox). JMAC emphasizes that speech events are player-driven rather than responsive to the sonic environment, in that speech events will be arranged in a waiting list to be cued by situational player actions. The speech AI then prioritizes the samples according to specific gamestates and pre-programmed scenarios, ensuring that samples that have already been activated will be dropped to the bottom of the priority list. In this case, the AI makes informed decisions on which samples to play in response to a player’s actions, although JMAC states that interactive audio functions as more of a “logic-based system” than a thinking machine. The perceived awareness is constructed by encoding relevant speech information into tools that will have logic parameters to know which samples are relevant, and can be prioritized for pertinent feedback in a given state, or in response to user actions.

AZ objects to the term “aware” to describe the function of audio tools, but rather would say that audio follows a “logical tree” for playback. “The AI signals the playback module to play the appropriate sound for a given action” so that “sonic behaviour is altered to react and mimic the gamestate in real-time.” Conversely, GD feels that awareness is an appropriate way to describe how AI logic responds to gameplay. He marvels at their ability, as sound artists at EA, to now create “sonic shifts” that are inserted at appropriate intervals in the game, according to priority sequences recalled by audio AI systems. He boasts that problems of masking don’t occur as they did in previous video game audio, because designers make decisions ahead of time to anticipate high-impact scenarios, and can program the AI in advance to decide which sounds will be heard over others, and which sounds will drop down in the mix without contributing to a noisy environment. “Cut-off events” are also integrated, says GD, so that any key sounds will not be lost in high-impact events. Speech comments pertinent to a specific event will be tagged with higher priorities for playback in game events, and “anything that is not played in the first event will get expired, so that it becomes a priority the next time a similar event occurs to avoid repetition.” All of this is controlled within speech tools. GD states that AI knows what will happen, and can play the samples precisely as events are happening. Furthermore, he excitedly comments on what he refers to as “clairvoyance”, which is when speech designers will actually program delays between game events and speech samples called “lag time” in order to adequately simulate the time it would take for an announcer to formulate an opinion and effectively comment

on it. The imposed delay will add to the realism for the player, avoiding what GD refers to as the “psychic announcer impression.”

As applied to interactive music, JMR suggests that it is more a system of prioritizing and feedback than actual awareness. According to JMR, interactive music responds to a player’s style and positioning within the sonic environment. “It’s a combination of what the player initiates, where the player is in the environment, and what the gamestate is. A specific action at a certain time in a gamestate will provoke a specific musical response.” Prioritizing is a challenge, but also a primary concern for interactive music designers, as the RAM must be programmed to manage up to 48 simultaneous voices at any one time, that must be balanced to limit masking, and any potential for a lo-fi environment in which key sounds could be drowned out in the mix. With less conviction than GD, he acknowledges that this is a challenge that is inherent in celebratory moments (when a goal is scored, for example) or reward sequences, because all of the audio components are generally programmed to sound off at once, all at the highest intensity to accentuate the sonic event. In order to do this, “duckers” (i.e. compressors, limiters) are implemented to stifle some audio components so that the key sounds integral to the meaning of the event will come through, and others will not compete for the player’s attention. For example, the music may fade down, or become a distant, ambient melody once a goal is scored in a sim game. Speech commentary may also be held back until after the crowd cheering subsides somewhat, thus creating a realistic impression of the announcers basking in the moment before interjecting. Event based music will commonly have parameters to either drop or augment the music, but may apply DSP effects as an alternative to volume parameters, such as reverb or filtering, to accentuate the music without increasing the amplitude within the overall mix. The trap, as JMR puts it, is predictability, which is planned against in the designing phase, to try and make interactive, variable music follow the experience without being intrusive. The emphasis is to incorporate “canned music”, but to “mix up” the music tracks and have music layers stored in RAM to have multiple tracks and samples that will play according to a particular theme or gamestate to avoid repetition. In discussing specific strategies incorporated to enhance interactive music in *Anarchy Online*, designer Bjorn Arve Lagim explains how musical variations are incorporated to create interactive combat music.

Each [music] layer can change into any other layer at any time... A typical sample has 3-5 transitions within its layer and a fair number of transitions to other layers... the layers function is used primarily in combat music, which changes to reflect how well –or badly- a player is doing in combat as well as illustrating the strength of the opponent. The combat music has 14 layers which can be activated and stopped at any time, depending on the situation. Mostly it reflects if the player is winning or losing, but it also reflects the size of an enemy.



If it's a huge monster, the battle music will be more intense, and if the opponent is tiny, the music will be less intense.<sup>214</sup>

Various input parameters can and are parameterized so that musical gamestates will “react on the fly.” JMR acknowledges that interactive music does not incorporate high levels of synthesized music construction at this point, and because of this, interactivity must be limited to some extent in musical passages in order to maintain manageability. He states “[sound artists] can really sculpt the sonic environment in [modern] games, but it's hard to keep in check if [designers] go too far with interactivity.

GD explains that prioritizing and balancing the overall mix is approached in a similar manner to the crafting of interactive speech, music and sound effect components. In addition to having AI for each of these auditory components, there is an audio AI that controls the overall audio construct to achieve a balanced soundscape. For MVP Baseball 2003, for example, GD explains that speech samples will be more drawn out to flow with the pacing of a baseball game. This functions accurately in game delays or drawn-out sequences because the audio AI is written and arranged with scripts and spreadsheets to create an audio AI event structure that is, as GD puts it, “hand-sewn”:

In the case of a long, boring sequence, there will be more chat time for announcers to talk about the weather and player stats between sequences programmed into the AI, as well as random crowd sounds to give an impression that they are growing impatient. If the player takes too long, hecklers will begin yelling in the crowd, hot dog vendors will call out more often, PA organ music will play; all of this will happen at once to simulate what would happen in the real baseball world... [The next time a game delay occurs] the AI knows how to play back everything related to how the same event might change depending on the context of [each event]. It's really a complex building process of programming AI tools that sound artists [need to experiment with] through trial and error to get the right playback modelling.

GD concludes this discussion by saying that sound design was previously generic and derivative, but that advances in audio technology are allowing artists to be more creative and expressive. The tools are catering to artists, allowing them to delve deeper interactively, and flex their creative muscles. Ultimately, GD says “the task is to relate audio to the emotional impact of the game, whether [the primary focus is] a detailed storyline or a complex atmosphere, depending on the game, to create a completely interactive environment” for players to submerge themselves in.

The sound artists interviewed spoke about the degree of audio interactivity that will depend on the genre of the game. IM explains that levels of interactivity will vary from game to

---

<sup>214</sup> Lagim, Bjorn Arve. “The Music of Anarchy Online: Creating Music for MMOGs.” Gamasutra.com, 2002. [http://www.gamasutra.com/resource\\_guide/20020916/lagim\\_01](http://www.gamasutra.com/resource_guide/20020916/lagim_01).

game depending on the goals of the game. Music in sim games, for example, does not need to be interactive other than to integrate short music bursts or jingles coming from the virtual PA system as would transpire in a real-life sporting venue. IM maintains that a high level of interactivity should generally be maintained, but emphasis on specific, sonic elements made to respond to the player will vary depending on its effectiveness in the game, and will also depend on the target audience. Fantasy sports games, for example, focus more on interactive speech and music elements, while sports sim games will focus primarily on sound effects and variable crowd ambience as the primary interactive component. IM remarks that performance is sometimes more of a priority than interactivity for audio designers. Having speech that has the appropriate tonal inflection (or paralinguistic variables) to reflect the emotion of the speakers is integral, as well as having sound effects that are vivid and dynamic, and music that creates or maintains the mood and pacing of the intended gamestate or event structure. Crowd streams in sim games are constantly being modified to correspond with cultural and geographic variables to enhance the realism (which is now culturally-specific, with designers collecting ambient soundscape audio from venues worldwide). Vocal talent focuses on getting the most out of the performance, which is made possible by hiring experienced celebrity actors and real-life sports commentators speaking the lines, and often ad-libbing the dialogue to create spontaneous and high-energy speech samples. Music is composed by professional musicians specifically for the game, or is integrated by designers, interactively, from the master tapes of licensed, popular songs. IM speaks of the complex art of crafting interactive audio, because it must be able to apply differently to every different context in the game. He compares it to the art of scoring for films, in that the music will be specific to its genre (i.e. music for drama, comedy, cartoon, thriller, etc.). The same applies to the video game industry, says IM, “there are intense, strategic simulation games, and there are fun, light-hearted fantasy games. Both need to incorporate high levels of audio interactivity, and the techniques need to be appropriate to the style of each.” IM maintains that the goal must never be forgotten: that the purpose of interactive audio is to enhance the fun, and make gameplay engaging. Even if one game is considered more interactive than another with regard to its audio adaptability, if it works for the game it is intended for, then it is a success.

JMAC also emphasizes the importance of paralinguistic performance for speech as being the main focus. “The emotion in their voice has to be believable, and achieving the finesse and realism in speech commentary [intonations] is the focus.” JMAC muses about previous speech in sims compared to modern games, acknowledging that the focus was originally just to have speech in games, no matter how generic and robotic. He laughingly refers to previous examples of game speech as being likened to a parrot talking, incorporating very little paralinguistic variation, and

often lacking any interactive consideration. The idea, as JMAC puts it, is to have speech behave in a believable manner. Having speech provide feedback to the player is essential, but first and foremost, the way things are said must be convincing. His goal in designing speech commentary for sports sim games is to minimize choppy transitions by constantly enhancing and updating AI speech logic to make real-time, informed decisions. The technique of “stitching”, which JMAC explains was the previously-adopted technique of suturing one-word samples together, is now avoided because it sounds choppy and unnatural, disturbing the natural flow of tonal speech patterns. Only full sentences are now incorporated to preserve the fluid dialogue created by the commentary. Levels of interactivity are crafted to suit the type of game. In sim games, JMAC explains, players expect an authentic, sports experience, which must be replicated to the finest detail through the audio. Voices heard in the crowd must therefore imitate the atmosphere, with specific comments, chants and cheers specific to localized crowds in MVP Baseball 2003. For fantasy sports titles like SSX Tricky, however, the player’s expectations are to “have exaggerated speech, with characters screaming over-the-top [exclamations], and to have Rahzel hamming it up with jokes and [witty] phrases.” JMAC maintains that the speech all functions interactively in that it provides suitable feedback for the player, but it must match the sonic expectations for the specific game genre.

JMR claims that interactive music is sometimes contained rather than used to its full potential in certain games, as he suggests too much interactivity would be counter-productive. For instance, music inherently needs to be a background element rather than serving as a foreground auditory signal for some games. If music were to shift with every move a player makes, it might detract from the auditory experience, and would become chaotic. Music that generates itself in real-time must be malleable, according to JMR, “the game audio should be so closely intertwined that you don’t distinguish one [sonic element] from another.” He explains that interactive music is all about making the interactive experience happen, but without overcompensating.

If the audio calls [unnecessary] attention to itself, then it has failed... [the goal of interactive music is] to shift the bio-rhythm of the game and move [players] emotionally through interactive, musical shifts... [these] are built on interactive audio moments that have shifts that stand out in the melee of video game audio...it should be felt, not recognized.

JMR and AZ both acknowledge that speech, music and SFX must function together as a single thread to incorporate all of the interactive audio elements simultaneously in audio gamestates. AZ says that the lines between sonic elements are necessarily blurred in designing the sonic environment to create an interactive experience, through which the rudiments of speech, music and SFX all function to position the player using the audio mix. He states that the audio should be

ever-changing, and that “speech, music and SFX should not, and cannot be segregated in order to achieve an interactive video game audio environment. Designers must work together to integrate all components, and be aware of all sonic elements.” In a bold statement, AZ states “The audio must function as a living entity, or we have failed.”

GD makes a point of stating that audio cannot be totally random or completely open-ended. “The audio should be interactive, but it also has to have rules.” GD describes techniques of integrating audio interactivity as “composed art” that designers must build in order to serve as a fun, interactive encounter for the player. He says that there may be drawbacks to implementing too much interactivity, in that players still want some degree of passive entertainment, and that designers have to know where to draw the line with interactivity. “It would be like giving someone a blank book and saying ‘write a novel’...players want a preset audio design; they want a world that has been created for them, and to see what it will do in relation to them.” They want to “build their own experience,” and control their own destiny in the game by “navigating through the simulated world.” GD points out how ridiculous it would be to incorporate “100% interactivity”, because that would be like giving players a sound-editing/sequencing program and putting them to work. Interactive audio must provide an element of control, but remain a fun, guided experience. Designers will always have a role in crafting interactive audio to position players in the auditory gaming experience, GD says.

## **4.2 Achieving Audio Interactivity**

The interviews conducted with sound artists at E.A. served to enlighten me as to how interactivity is applied by the professionals who design and implement the audio inherent in modern games, and to get at the root of interactive considerations in video game audio. The comments made by designers confirmed my assumptions that audio interactivity will vary depending on the game and the genre for which it is intended. Their statements also verified that audio engines, while not necessarily “aware” of a player’s actions, are programmed to adequately respond in real-time to player actions and style variations.

Before conducting the interviews I was unaware of the painstaking efforts that are made by designers to program interactive audio tools for every possible sonic variation in order for real-time audio AI to react and adapt to the player’s actions. While I was conscious of their need to anticipate scenarios to create a realistic and immersive sonic environment, I was previously oblivious to the intricate detailing, composition and arrangement that must be conducted by a team of sound artists in order to program tools and audio AI in order to achieve audio that ebbs and flows with gameplay. While it is clear that audio does adapt itself to changing gamestates and

environmental aspects in the game, the interactive nature of real-time audio technology surpasses auditory feedback principles, serving not only as an aural representation of the graphical components, but signalling information beyond what the visuals are able to portray, and intended to play on the user's emotions in order to immerse the user into the surrogate environment. Subtlety through understated auditory cues, unpredictable shifts in audio components in relation to player actions, and perceived awareness and intelligence of virtual components in the game brought about by adaptive audio advances have elevated video game sound onto a similar plane to that of virtual graphics. After conducting analyses and having interviewed professional sound artists, I would suggest that audio may even have more of a role in potentially immersing the player into the interactive gaming experience than visuals, in that audio is designed to function on a more subconscious level to provide emotional fluctuations during gameplay.

IM repeatedly emphasizes the importance of layering audio streams in order to simulate audio awareness through real-time audio that "reacts to situations in the game instigated by the player." IM alludes to the fact that interactive audio algorithms may not necessarily be aware of a player's actions, but that sound artists, through techniques of mapping, scaling, and "parameterizing" sonic events, program audio AI to be able to prioritize appropriate sounds to be triggered in the apposite game scenario in real time. During the interview with IM, I probed whether the audio process is truly an interactive one, in that it can be argued that players do not have full control over the audio, but rather follow a predetermined path of audio events that have been programmed ahead of time into the AI. IM maintains that, while scenarios are pre-programmed and arranged by sound artists, users still have control over the gaming process, and while players may not directly control the sound, they control the game action which in turn triggers the appropriate sonic scenario. "The game action controls the sound, the player controls the game action" he states. IM does not see indirect control of audio by the player as being an impediment to interactivity, and sees this distinction as a moot point, in that the potential for an interactive relationship is still inherent between the player and the sounds that are played back in the game.

JMO is optimistic that powerful processors devoted exclusively to sound will one day allow sound artists to explore their full, creative potential for interactive audio without CPU memory or RAM limitations. He insists that more random variables programmed with input parameters into music AI will give the impression of a variable soundtrack that intends to create a personalized, individual interactive entertainment experience.

JMAC praises developments in audio interactivity, claiming game audio is aware and able to provide suitable feedback to form a real-time dialogue with the player. He marvels at the

advances in audio AI and at the increase in audio creativity inherent in games with sound artists at the helm, maintaining control over their craft. Much like IM, JMAC highlights the need to focus primarily on the arrangement and layering of sounds to create an ever-changing sonic environment in relation to the player, what he refers to as “stream management.” JMAC foresees increasing creativity and interactivity in game audio as more resources will be at the disposal of sound artists, allowing them to further enhance the emotive impact and atmosphere provided through audio, which he believes will surpass that of Hollywood films. He foreshadows many exciting advances in interactive audio that will transpire as video games continue to be recognized as “interactive entertainment” for a variety of genres spanning different target audiences.

Audio Leads JMR and GD both allude to more subtlety in audio cues, and a shift away from previous generic-based sound models. GD is optimistic that an increase in methods of applying synthesis, sample-shifting and granular techniques will model game audio much like real world soundscapes; reactive to their environment rather than generative audio. GD offers an interesting perspective when he stresses the need for order and structure in game audio, and that there should necessarily be limits to the amount of interactivity incorporated. He refers to the design of interactive soundscapes as ‘composed art,’ that cannot be totally random or open-ended. He states “you need rules in an operational world; there must be goals... a story to explore.”

AZ stresses the need to perfect the interactive story in order for the player to perceive change throughout the narrative. While the sonic environment functions to immerse the player, there is still a need for “interactive audio to reinvent itself” in order to evolve throughout the interactive storyline. AZ emphasizes the need for audio design to do more than incorporate a basic evolution of amplitude, pitch and timbre. He stresses the need for more designers trained specifically in the “interactive aesthetics of sound” to craft an interactive story that “has the ability to be ever-changing through the audio.” While we are not there yet, AZ says audio is well on its way to incorporating memory storage for the game to remember what the player’s style is, and to know what sounds have previously been played so that it can alter future playback. AZ also maintains that the lines between speech, music and SFX are necessarily blurred in interactive sonic environments in order to adequately achieve the interactive experience sought by designers. The overall mix, AZ states, must be composed of sonic elements to form a single thread that incorporates all of the interactive elements simultaneously in audio gamestates. He stresses the need for audio designers to work together as a team, to integrate all components, thereby forming an interactive sonic environment that functions as a living entity, intended to immerse the player.

When applying the goals of sound artists to the audio environments inherent in the games deconstructed in this study, it is unmistakable that the efforts of sound artists are designed to allow for more interactivity. Fewer technological limitations and more control for sound artists over the design process has allowed for audio designers to craft their auditory visions, and to cater the interactive audio experience to players without being overshadowed by graphical considerations. The designers note that there are still some technical hurdles to overcome before CPU memory, audio RAM, Dolby 5.1 Digital audio, optimal disc speed and uncompressed, high-resolution samples become standardized and commonplace. As AZ asserts, audio technology is rapidly advancing, but sound artists must cater to the lowest common denominator, assuming that the majority of gamers will be playing video games on a stereo television with poor speakers, and that many of the sophisticated considerations in the overall audio mix may be lost on players using a standard TV receiver. Once televisions capable of reproducing Dolby Digital surround sound become commonplace, AZ says, sound artists will have free reign to push the sonic envelope into that domain. Although the video game industry is now a lucrative one, financial constraints and limited budgets still prevent sound artists from having the freedom to fully explore and apply their auditory visions. However, the hurdles are becoming fewer now that interactive audio has garnered some much deserved attention. The consensus of the designers interviewed was that technical constraints will become a non-issue with advances in future console gaming systems such as Sony's *Playstation 3*, and that higher resolution and more RAM devoted to audio will allow designers the freedom to focus more on interactive considerations. While he acknowledges that there is still work to be done to overcome some of the technological limitations, IM feels they are becoming (and will continue to be) less and less of an impediment to sound artists as more resources are devoted to game audio. IM emphasizes that algorithms need to be continually re-analyzed to incorporate more complex layering so that separate audio streams work together harmoniously to form a seamless audio environment that is designed to be fully flexible and responsive to the player.

In comparing the statements made by the designers to my own analyses, I can conclude that randomness is indeed achieved, made possible by unpredictable audio components that insert themselves appropriately in real-time to coincide with the player's actions. Subtle cues in speech, music and sound effects are intended to evoke responses from the player in such a way that players can feel changes in the gaming atmosphere, and can conduct themselves accordingly, potentially altering their play style to perform properly within context of the gaming environment. Sophisticated music that is composed and designed to function interactively through real-time algorithms, remixing the score in relation to the player's style of play as in *SSX Tricky*,

demonstrates that advances in real-time scoring have the potential to adapt to the player, designed for the purposes of creating a different musical experience each time the game is played. Subtle shifts in the score depending on the player's performance are designed to cue players without blatantly instructing them, intending to add to the believability of the gaming experience without calling attention to the artificial nature of the virtual context. Variable SFX, such as the complex crowd sound inherent in sim games like *MVP Baseball 2003*, consist of multiple loops and layers that adapt themselves in real-time to game events, each layer having the capacity to adjust according to any number of input parameters. The intended result is a crowd ambience that is ever-changing and seamless, made to enhance the believability of the simulated sports experience, thus creating a potentially realistic environment to immerse the player into a virtual sports arena that can respond to the player's performance. Unpredictable, reactive speech AI that takes instruction from the player as in *SOCOM: U.S. Navy SEALs* introduces speech recognition in game audio that has the capacity to create an ongoing dialogue with the player throughout gameplay. JMAC is optimistic that interactive speech will continue to advance in the near future. With more attention focussed on speech performance and fewer instances of 'stitching' in order to preserve the natural flow of speech patterns, a realistic dialogue is formed to guide the player through gameplay in real time. Paralinguistic performance as a primary focus serves to heighten the realism of the experience, and the act of speaking to virtual characters and having them respond accordingly with pertinent comments serves to authenticate the participatory experience. The designers' intended experiences are crafted to be immersive, realistic gaming environments that serve to provide suitable atmosphere and emotional submersion, thereby surpassing the previous experience of playing a video game, and intending to transcend the participatory activity into an interactive entertainment experience. Before conducting my own analyses to deconstruct the virtual sonic environments within current game titles for different genres, let us first highlight some of the designing principles, issues and viewpoints from some of the leading sound designers in the interactive entertainment industry today, working at Electronic Arts Inc., "the world's leading independent developer and publisher of interactive entertainment software."<sup>215</sup>

---

<sup>215</sup> Electronic Arts Inc., "About EA" EA.com, 2003.  
<http://www.ea.com/global/corporate/aboutea.jsp?src=11001homecorpabounone>.



### 4.3 Applying Interactive Design Principles

Having conducted the interviews with designers, I can now apply their principles of design for implementing interactive audio considerations to the case study to be conducted in the subsequent chapter. The feedback from the sound artists documented in this chapter, while distinct from the research material applied earlier this study, still serves to guide my analyses in extracting the different approaches to interactive audio design adopted by professional sound artists. A typology of the different design principles obtained during the course of the interviews can now be condensed.

Different levels of interactivity achieved through video game audio will vary depending on the game genre for which they are intended. Some games will require an emphasis on spatial characteristics in order to create realistic sonic environments, through which the player has the freedom to explore, as in first-person shooter games, graphic adventure games and role-playing games. The focus in the above mentioned games requires a vivid construction of the auditory elements within the sonic environments in efforts to immerse the player within the surrogate gaming context. Other game genres, such as sports game simulators, fantasy sports games and music simulation games will rely on strategies for incorporating a high level of audio interactivity for the sake of providing functional auditory gamestates and game events. This requires designers to incorporate logic and variable input parameters that can be altered in real-time by audio algorithms, reactive to the player's style and actions initiated during gameplay. The role of interactive speech is to provide feedback for the player through linguistic and paralinguistic emphasis, whether subtle or blatant, to provide information to the player about the user's adopted style of play. Music will follow conventional, predetermined song structure for some game genres, while other games will require interactive rises and falls to develop the score in relation to the player's character, thereby adapting to the mood of the experience to create dynamic musical shifts. Sound effects are designed to be both logical and unpredictable, made to defy user expectations, providing elements of surprise and unexpected outcomes, while delivering aural cues to indicate whether the player is taking the appropriate approach within the game. Through parameterizing and prioritizing samples to correspond with gameplay initiated by individual players, the strategies and goals of audio designers are to make sounds appropriate to the genre, and to make the gaming experience different each time the game is played, thereby augmenting a given game's replay value.

Perceived "randomness", or unpredictability of auditory events is strategically implemented to enhance the believability of the surrogate auditory context. Layering is a tactic that is adopted by sound artists to make audio components react appropriately to game events.

Interactive speech must flow naturally and repetition must be avoided at all costs to maintain the believability and perceived intelligence that is sought after, driven by a player's actions. The method of "intervening samples", described as cycling groups of sounds over a number of successive games, is a strategy incorporated to avoid repetition, thus increasing believability. For music, unpredictability is achieved through real-time layering of instrumental components by algorithms that are designed to shift the music to flow seamlessly along with individual gaming experiences. Interactive music functions best when it is written for the purposes of being sculpted in real-time, catering to individual styles of play. The layering of music samples in real-time is intended to create a soundtrack that ebbs and flows in relation to each player's evolving performance for a personalized, interactive experience. Sound effects are designed through digital signal processing (DSP) parameters in order to create an ever-changing perception, incorporating logic data in SFX AI to trigger appropriate sounds to correspond with different gamestates.

Audio cues are implemented in order to signal players to take specific actions. These can also be subtle or obvious depending on the genre and gamestate. Instructional sequences will incorporate more blatant cues for the player, while more advanced in-game sequences will require subtle hints in order to signal changes in the environment. Speech commentary, for example, is often designed to signal something indirectly to the player through paralinguistic inflections rather than linguistic variables to direct the player. Designers rely on the scripting of more subtle speech cues to guide the user through insinuations rather than instructing the player. Musical shifts and sound effects are also designed as subtle cues to provide tension and atmosphere through the soundtrack. Interactive music and SFX are intended to play on emotions rather than logistics, designed to make the player *feel* a change in the mood of a particular gamestate. Techniques of filtering and DSP modulations, for example, can give subtle indications about shifts in the virtual sonic environment.

Variable input parameters encoded into speech, music and SFX AI serve to make audio reactive and responsive to the player in real time. While the term "awareness" may not adequately describe the logic process in designing AI to vary audio components in real time, it is evident that the audio tools are logic-based systems that are programmed to prioritize samples to follow a plausible progression of auditory sequences. In designing interactive speech and SFX AI to play the appropriate sounds for any given scenario, sonic shifts are programmed to be inserted at appropriate intervals according to priority sequences that are recalled by audio AI systems, thereby avoiding repetition, and enhancing the believability of the virtual sonic environment. In "high-impact" scenarios (when multiple voices are activated simultaneously) "cut-off events" are integrated by designers to ensure that key sonic elements will not be lost or overshadowed during

game events. "Tagging" sounds according to a priority scale allows for the AI systems to recall the proper sounds and ensure they stand out as signals within the context of the virtual sonic environment, achieving a balance, thus maintaining a hi-fi soundscape. Music is also incorporated using priority strategies so that a specific action at a precise interval in the game will provoke the desired musical response. Input variables are "parameterized" so that music will conform with gamestates in real time. As with cut-off events in the case of speech and SFX, "duckers" are implemented in music AI systems to compress or limit certain musical components so that the melody will either fade down or become a foreground element, depending on the event being designed for.

While the sound artists interviewed maintain that a high level of interactivity in the overall audio construct is what is sought after, the degree of reactivity to the player will vary depending on its effectiveness in the specific game. Performance of the audio component as a whole is key in making the audio design effective for the purposes of immersing the player within the fantasy environment. The art of designing sound for games can be likened to scoring for films in that the techniques for crafting the soundtrack will depend on the genre for which it is intended. Principally, the challenge for designers is to relate the audio components to follow the emotional impact of the game being designed, and to create sonic environments with variable degrees of interactivity depending on the genre. Designers must be concerned with programming audio algorithms and AI to make logical, informed decisions in real time to match the expectations of the game genre. There is sometimes a danger in incorporating too much interactivity, in that an over-abundance of dynamic shifts might detract from the experience, potentially resulting in a noisy, chaotic, lo-fi environment. Interactive speech, music and SFX must function as a whole, designed to position the player within an immersive auditory context. Each of the designers interviewed stress the importance of teamwork and collaboration of an audio design team to ensure that all of the interactive audio components work cohesively to function as a living entity, intended to create an engaging, participatory experience for each individual player. Having conducted the interviews with designers, I can now apply the design principles and strategies documented in this chapter to my own video-game analyses in the form of the case study to follow.

*...from simple button clicks on a menu screen to intricate character movements, weapons, environments, background ambience, and so on... providing this type of feedback to a player is, in essence, the magic that draws them into the virtual world. They interact with the game and are rewarded with aural and visual cues adding to the feeling of being inside the game making things happen. Sound [...] play(s) a grand part in this design and, at the onset, the developer will want to cover every possibility... The sounds should all work together to enhance gameplay, not aggressively compete. The last thing we want to do is cause the player to turn the sound off.<sup>216</sup>*

-Aaron Marks  
Video Game Audio Designer

## **5. Case Study: Video Game Textual Analyses**

This chapter applies an experimental design to analyze audio components of specific games to document the sounds exclusive to each game; to examine and evaluate how the sounds function within the context of the virtual sonic environment; and finally to hypothesize about the potential for audio interactivity within each game.

### **5.1 Methodological Framework**

The analytical framework to be used for the case study in deconstructing game audio for three specific games in this chapter will be based on the *speech-music-soundscape* continuum in order to investigate sonic elements within the surrogate audio environments. The sounds will first be observed as separate entities, categorized as elements of speech, music or sound effects in order to determine the function of these sounds within the soundscape of each game. By examining elements of music, speech, and sound effects within the soundscape as elements that make up the environmental audio context, and scrutinizing how the auditory aspects interact within the overall electroacoustic design, I can better hypothesize about interactive audio elements within the games I am studying.

#### **5.1.1 The Method: Playing By Ear**

Using information conveyed in the previous four chapters, I will be incorporating knowledge and information specifically pertaining to video game audio, interactivity, and using acoustic, soundscape theories, and feedback from designers to deconstruct the surrogate audio environments designed for the games in question. Through the analytical process, I will be

---

<sup>216</sup> Marks, Aaron. [The Complete Guide to Game Audio](#). CMP Books, Kansas: 2001. p.246-247

isolating the areas of speech, music, and soundscape respectively, in order to evaluate the juxtaposition of sonic aspects, and hypothesize about their overall interactive potential for the player. It is important to note that this is somewhat of an unconventional approach to analyzing video game audio (although there is no single approach within this new area of study that can be considered conventional at this stage), as the lines between the three systems are typically blurred, and are considered together (and often inseparable) in order to compile interactive audio elements effectively for the entire audio spectrum of the game. The process of incorporating interactive audio, as was explored in detail when discussing the principles of audio design, which was the focus of chapter four, incorporating interviews conducted with current sound designers at Electronic Arts.

Analytical categories within each system of the continuum will be similar, but will apply to separate audio components, as they are different when pertaining to speech, music, and sound effects, the latter including environmental signals, ambience, soundmarks and keynote sounds. I will be looking at each game separately to first pinpoint the sounds themselves before determining their function within the game audio context as a surrogate soundscape, and finally analyzing its interactive audio construct, contrasting it with that of other games from contrasting genres. First I will examine elements specific to each system within each game: that is, linguistic and paralinguistic aspects for Speech; melodic, timbral, rhythmic and harmonic elements for Music; timbre and dynamics of sound effects for ambient sounds, etc. I will be looking at acoustic parameters of some key sounds present in the soundscape of each game to determine their relationship to one another as signals, keynotes, and ambient sounds within the virtual environment, as well as localization and listener/player positioning. Then psychoacoustic perceptions will be applied in terms of potential moods and semantic meanings that are implied and responses that are intended to be evoked through the elements of speech, music, and soundscape. Interactive components and the potential for user participation specifically focusing on audio elements such as sound indicators, informative agents and cues (such as soundmark signifiers, reward music, etc.) will be examined for the sake of interactive immersion. Hypotheses about the potential for perceived realism within the virtual gaming context will also be a component of the analyses. Negative drawbacks and issues arising from game audio will be saved for consideration in the chapter to follow. Once I have conducted textual analyses of video game audio in the following cases, and having documented the principles of interactive audio design as articulated by sound designers in chapter four, I can then proceed to put forward my own conclusions about the overall interactive audio construct with regard to immersing the player within the participatory gaming experience.

### 5.1.2 Chosen Games, Genres and their Significance

The games to be analyzed as video game audio "texts" in this case study are *SOCOM: U.S. Navy Seals* (Sony Computer Entertainment America, 2002), *SSX Tricky* (Electronic Arts Inc., 2001) and *MVP Baseball 2003* (Electronic Arts Inc., 2003) for the PlayStation 2 home gaming console system. The games were chosen for their unique and distinctive treatments of audio, featuring some of the most advanced interactive sound components to date. After researching articles about video game audio in web-zines, observing and sampling the gamut of the newest games, and interviewing sound artists at Electronic Arts, I concluded that these three games exemplify some of the key technological aspects of interactive audio that I have articulated in the previous chapters of this thesis. The games all epitomize the interactive audio design considerations put forth in this dissertation, but were also chosen because each one can be contrasted as emphasizing a particular interactive element within the continuum. Each game has also been chosen as a representative of contrasting game genres to illustrate how interactive audio can vary, spanning diverse game classifications. These games are at the forefront of interactive video game entertainment as some of the most popular being sold and rented by gamers today. They have been selected for this study because they embody some of the key concepts I have highlighted as elements contributing to audio interactivity in video games. In preparation for these analyses I have deconstructed and catalogued the key sounds that make up the surrogate audio environment that encompasses the player within the fantasy world context of each game (Appendices B,C,D). I have also formulated a detailed chart to illustrate the sonic elements in relation to each other within the functional soundscape as encountered in sequential order during gameplay of Game 1 (Appendix A). Each game analysis will consist of an examination of the game's discrete sonic components, its functional environmental context, and its potential for interactivity.

### 5.2 Game 1: SOCOM: U.S. Navy Seals

*SOCOM: U.S. Navy Seals* (SOCOM), is an example of the popular First-Person Shooter games. SOCOM takes the FPS genre to new levels with 7 difficulty levels, 12 missions in 4 distinct global locations (Alaska, Thailand, Congo and Turkmenistan), and provides for intense team-based gameplay in "the most authentic military combat simulation" to date.<sup>217</sup> The game attempts to portray realistic U.S. Navy strategy, having collaborated with the U.S. Naval Special

---

<sup>217</sup> Gamespy, 2002. <http://www.gamespy.com/SOCOM/inside/gamefeatures.asp>.

Warfare Command to adequately simulate the interactive gaming experience.<sup>218</sup> SOCOM engages the player from the initial menu screen, with an actual U.S. Navy Seal documentary as one of the game's features to make the player feel as a part of a real, justifiable military cause. The documentary rouses the player into battle, glorifying teamwork and the role of the U.S. Navy. Although SOCOM is a game that requires the player to shoot enemy characters as a strategy to complete missions, it is important to note that SOCOM takes a unique approach within the FPS genre. The objectives require the player to explore levels and gather enemy intelligence in order to deconstruct terrorist plots, and necessitates that the player approach missions surreptitiously, and adopt a covert, tactical attitude. Rather than simple objectives of shooting to kill (as is the single purpose of most FPS games), SOCOM requires that the player incorporate teamwork and adopt strategic approaches and stealthy, non-violent tactics to complete detailed, multi-purpose missions, which involve capturing and subduing enemies rather than killing them. Mission objectives include ambushes, hostage rescue, bomb diffusion, demolitions, laser designating targets, reconnaissance, eavesdropping on terrorist conversation, clearing buildings, etc.<sup>219</sup> The game can be seen as promoting alternative strategies for problem-solving instead of annihilating everything in sight. Players are actually discouraged from the act of killing, and often times will be criticized or punished with setbacks in the game's progression for taking violent action. The "HQ" (Navy Headquarters) voices will scold the player for blowing their cover and for compromising the mission, and in extreme cases, gameplay will end with the words "Mission Failure" in bold letters across the screen, and the player will be forced to replay the level, adopting more premeditated and less aggressive strategies. An example of this can be seen in stage 3 entitled "Oil Platform Takedown," which requires the player to defuse bombs, undetected. If the player fires a gun and is detected, the enemies will alert their leader that they have been discovered, and the mission will end abruptly. SOCOM also offers a variety of modes for online gameplay to propel the gamer into a virtual world that spans the worldwide SOCOM community in the Internet. The online component allows for 16 players from across the globe to engage in team play simultaneously through a broadband connection rather than being confined to single-player mode on the home console gaming system.

---

<sup>218</sup> Ibid.

<sup>219</sup> Ibid.

### 5.2.1 SOCOM: Speech, Music & SFX

I have catalogued the sounds I perceived while analyzing the first 3 stages of SOCOM (Appendix B), as the signals, keynotes and ambient sounds that contributed to the overall sonic environment. Music and sound effects play off one another in such a way as to enhance the suspenseful nature of the mood intended for the gaming context, while the speech component serves to guide the player through the participatory experience.

The speech component in SOCOM is perhaps the most impressive aural constituent. The sound design team at Zipper Interactive, Inc. implemented a special peripheral device specifically for use of this game on the PS2 to enhance the auditory experience and immerse the player. A USB headset with microphone can be purchased along with the game to allow the player to enjoy the full audio environmental experience. The instructions and tactical strategies from the "HQ" (Navy headquarters) voices are spoken through the headset, as are the voices of the other three members of the SEAL team fighting along-side the first player, who exists in the virtual form of "Kahuna," the Navy SEAL team leader. In online gameplay the headset allows for voice-chat so that the players can communicate to one another in real-time while playing, using a broadband Internet connection. The headset simulates the voices as they would be heard transmitted through a walkie-talkie, and adds to the realism of the gaming experience, as the headset and microphone are used throughout gameplay for the player and the digital teammates to communicate messages to one another to foil terrorist plots. The method of transmitting the "HQ" instructional commands along with the voices of the player's SEAL teammates through the headset allows for a more balanced perception of sonic elements. The allied speech components are mostly reserved for the headset, which creates a greater distinction from enemy voices and sound effects emitting from the television speakers. Although the quality of the voices sound "tinny" and the spectral range is of a limited bandwidth, the voices are clearly enunciated, and are generally varied in tonal colour to convey semantic meaning in expressing excitement, urgency or caution. The peripheral headset eliminates the potential problem of competing sounds or masking, in that the voice commands of the player's allies can be clearly heard without being drowned out by the music or sound effects. Consequently, music and ambient sounds do not need to be dulled down to ensure that speech components are audible in the foreground.

Next, the music is made up of a primary orchestral composition that contains multiple variations to maintain continuity throughout the game. Subsequent melodies are all variants of the main theme. The original score, composed by Jeremy and Julian Soule of Artistry Entertainment Inc., provides the perfect musical atmosphere for the gaming context with triumphant fanfares, suspenseful movements, tragic ballads and overtures that glorify the player. The musical excerpts



used in the game correspond to gamestates, either representing victory and confidence with a fanfare in a major key with harmonic overlays, or symbolizing tragedy and loss in the musical form of severe-sounding renditions in a minor key, with inharmonic, discordant juxtapositions. Suspense is conveyed through inharmonious drones, irregular drumfills treated with reverb, rests, and sudden jarring bursts to keep the player on edge. The music for *SOCOM* exemplifies the timbral sophistication of orchestral scoring in modern games, with dynamic musical excerpts that use a wide range of variable tones and textures to create gestural movement and progressive shifts throughout gameplay. The quality of the music demonstrates the advances in audio resolution, sending out clear, digital, stereophonic musical samples that the player can appreciate for their clarity and precision.

Sound effects can be seen (or heard) as paralleled with musical components, in that the sound effects consist of high-quality, vivid and dynamic samples that enhance the realism and immersive nature of the gaming experience. Gunshots, explosions, and helicopter sounds, for example, were collected on-site with trained Navy Seals using U.S. military-issue weaponry to maintain the highest level of realism.<sup>220</sup> The quality and clarity of the gunshots are vivid, jarring, high-impact bursts, coupled with DSP manipulations and compression that serve to enhance the impact of the sounds, and sharpen the attack of the gunshot sound envelope. Silenced weapons are muted appropriately, but still provide impact with the "whizzing" sound of bullets hurtling past the player and ricocheting off the walls and surrounding surfaces. Ambient, environmental sounds, such as the intermittent howling of wolves and screeching of crows in level 2, enhance the realism of the external context, as well as contributing to the eerie mood established by the music. Sound artists implemented newer methods of sound manipulation to enhance the realism in *SOCOM* to simulate actual acoustic occurrences. Most impressive is the simulation of a Temporary Threshold Shift (TTS) that occurs in the game when a grenade explodes in close proximity to the player. A large threshold shift, Barry Truax explains, is "brought about by a saturation of the hair cells" within the cochlea, which acts to protect the auditory system against loud sounds by providing a temporary "period of aural 'rest,' in order for the hair cells to be rejuvenated."<sup>221</sup> Instead of the explosion sounding with a sharp attack and almost instantaneous decay as in most action games, the sound effects designers took into account the dampening effect of the aural rest period that ensues subsequent to the ears being suddenly impacted by an excessively loud burst of sound. To simulate this effect in the game, the explosion sound is

---

<sup>220</sup> Sound artists can be seen collecting these sounds as part of the documentary footage included in the PS2 game's DVD footage. The documentary can be accessed by clicking on "Documentary" in the introductory menu screen of the *SOCOM: U.S. Navy Seals* game, Sony Computer Entertainment America, ©2002.

<sup>221</sup> Truax, Barry. *Acoustic Communication*. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.16

immediately followed by a fade-out, muffling and dampening all sounds heard over the next few seconds, coupled with a high-pitched squealing sound to represent tinnitus; the ringing in one's ears that occurs with TTS. Sounds gradually return to their normally-perceived volume level after a few seconds, replicating (though shortening) the time it takes for the ears to adjust before being able to perceive sound normally.

### 5.2.2 SOCOM: The Virtual Sonic Environment

Because of the covert nature promoted in SOCOM, the audio must also function to immerse the player in a suspenseful and tension-filled environment. SOCOM has been hailed as one of the best games for the PS2, and it has been glorified for its audio components. Ryan McCarthy in his review of SOCOM for *PS2 Central* pinpoints the musical score, sound effects and voice acting as being the highlights of the entire gaming experience.<sup>222</sup> McCarthy describes SOCOM as an "acoustic delight," marveling at the realistic sound, and inviting the reader (and potential player) to experience the game to "truly feel the war in your ears."<sup>223</sup> It was difficult to single out many of the sounds as separate entities, as they all function together to create the vivid environment specific to each scene. This is the desired effect for designers in making the sound components flow together seamlessly, and for the auditory experience to blend together instead of the player singling out specific signals in the mix.

The speech components function not only to guide the player through gameplay (with instructions from "HQ" and Kahuna's fellow teammates), but also to cue the player and enhance the realism of the intended U.S. Navy strategy simulation. Each voice has different spectral and timbral characteristics that distinguish one team character from another. The player's character is the commander "Able" figure, Kahuna, who utters grunts when shot, but has no spoken dialogue in the game (other than the commands issued by the player's voice, assuming the role of Kahuna in the game). The secondary "Able" character, or sidekick "Boomer," has a husky, low-pitched voice, while the "Bravo" characters, "Jester" and "Spectre," feature more high-pitched voices. The voice of Jester is sharp and high-pitched, and speaks in a southern accent, while Spectre's voice has a smooth tone, incorporating more mid-range frequencies. The voices of the teammates serve as recognizable keynote sounds for the player. The "HQ" female voice is graceful and pleasing to the ears, although stern when she issues warnings, serving as both a reliable and authoritative source of information. The male "HQ" voice is more subdued, calm, and states facts plainly, with little variation inflection, implying factual information without bias. The enemy voices, spoken in

---

<sup>222</sup> McCarthy, Ryan. "SOCOM: U.S. Navy Seals." PS2 Staff Review, 2002.  
<http://www.psx2central.com/reviews/psx2socom.htm>.

Russian, presume that the player will not understand the words linguistically, so the intonation or paralinguistic inflection will be exaggerated and pronounced to signal alarm, oblivion, anger, compliance or fear to the player. English translations are provided through subtitles, but the voices convey the meaning clearly, especially when the enemy characters utter a screaming war cry prior to attacking Kahuna from behind, which serves to cue the player that only seconds remain before being attacked. The enemy voices serve as a soundmark or symbol, in that a cautious attitude is adopted as soon as an enemy voice is heard. The piercing war cry of the enemies function psychologically as a sound phobia, startling the player, thus creating a sense of fear and alarm that an enemy is quickly approaching. These sounds are crucial in the overall effectiveness of the gaming context, as the character voices function independently from visual stimuli. Habitually the player must be cued to action solely by auditory cues, as they are the only indication that an enemy or ally is nearby, but not in view. In this sense the effectiveness of the visuals is sometimes dependent on audio in SOCOM to deliver the appropriate mood and response for the game to be effective.

The music during gameplay is not overpowering, but serves to enhance the mood of the gaming context, musically, and does not distract from gameplay. Cinematic non-interactive sequences (NIS) are integrated nicely to function as mini-movie segues, advancing the plot and allowing the player to relax and reflect on the recent victory or accomplishment. The musical accompaniment for introductory, closing and credit sequences function effectively according to Aaron Marks, to "give the player a sense of closure after the effort of playing the game," and that "the fanfare will reinforce the movement."<sup>224</sup> The music accompanying the menu screens in SOCOM functions effectively as well, maintaining the mood and energy level created in gameplay, but also providing a moment of rest for the gamer to prevent "over-stimulation" and serve as the "calm before the storm" in the transition to the next challenging level.<sup>225</sup> The music in SOCOM functions to keep the player immersed in the virtual sonic environment, and evolves along with the player's progressions through the storyline. While looping does occur, and crossfaded transitions of contrasting melodies are frequent, the music layout is done in an aesthetically tasteful manner that does not call attention to alterations, and the orchestral, non-rhythmic nature of the scoring does not necessitate metrical synchronization, or *beat-synching*. Reward music is incorporated in the form of a short, three-second fanfare, which sounds each time an objective is completed, as well as a triumphant, celebratory theme in the accompanying "Mission Successful" menu screen. This sample serves as a soundmark for the player that acts as

---

<sup>223</sup> Ibid.

<sup>224</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.188-189

a musical symbol, and creates identification with this excerpt. The player feels a sense of accomplishment and reward each time this sound is played. Conversely, when the player fails to accomplish a goal, the music in the menu screen following a mission failure will reflect the negative mood in a minor tone to signal defeat. Aaron Marks explains that these musical signifiers act as cues to the player, but must be treated delicately by audio designers to either signal victory while keeping the player focused on the challenges at hand, as well as balance impressions of encouragement and defeat without discouraging the player.

When you win, an optimistic flourish of sound rewards your superior efforts. When you happen to be less successful, the music is either demeaning or mildly encouraging, prodding you to try again... Winning cues will tend to be upbeat, composed in a major scale, with a lot of pomp and circumstance... capitaliz[ing] on the basic need for recognition... A loss can be presented to the player as a defeat or minor setback... Leaving them with a sense that they were still successful and can do better next time...<sup>226</sup>

SOCOM personifies each of the above elements in its musical design and structure. The music follows the experience to create the desired mood and atmosphere suitable to the environmental context and situational gamestate.

The sound effects serve to enhance the realism within the sonic environment. From the sounds of helicopters and wind, to snowmobile engines and swirling snow, to underwater bubbles and waves, to boat engines, fans, and radar beacons; the sound effects in SOCOM are highly vivid, and function to situate the listener within the surrogate environment. Lifelike ambient sounds, such as the wind and waves in level 1 and swirling snow in stages 2 and 3, provide contextual meaning and significance for the player in establishing the physical conditions of the environment. The timbre and transients of gunshots vary so that friendly fire (typically defaulted to a silenced rifle) can be clearly distinguished from enemy shots (commonly represented by blaring machine gun fire). The sharp attack of the enemy machine gun envelope often functions to startle, alert and alarm the player, while the softer muted blasts from the silenced rifles of the SEAL team serve to reassure the player that the team is in close proximity. An impressive sound effect that can be said to establish context and culture within the soundscape of the game is the radio in stage 2. In each abandoned cabin of the stage 2 ghost town setting that the first player character enters while searching for enemy files and information, the sounds of radio static and muffled music can be heard to signal to the player that an enemy guard is nearby. When Kahuna approaches the radio on the table, the sound of radio static amplifies, as do the muffled tones heard under the static crackling. As the player draws closer and the music gets louder, it becomes

---

<sup>225</sup> Ibid., p.189

<sup>226</sup> Ibid., p.191

clear that that music is actually a culturally-specific song, sung in the foreign enemy language (Russian, in the case of stages 1-3). Even more impressive is the attention to detail in that each radio that is encountered in the numerous cabins occupying the compound will each play a different Russian song, spanning diverse musical genres, to augment the realism, in that different characters would prefer different styles of music.

Spatial characteristics are implemented through sound effects, in that the soundscape will establish the player's position, and will give indications about proximity to sonic objects. For example, the mechanical buzzing, humming "whir" of the boat engine fans in stage 1 can be heard when approaching a doorway to the engine room containing the fans. One of the objectives in stage 1 is to deploy an explosive satchel in the engine room of the vessel, and the boat engine sound serves as a sound signal for the player that the room being sought after is nearby. If the player turns to face an alternate direction when in close proximity to the boat fans, stereophonic localization will be taken into account, as the context is directional, and will provide the appropriate aural simulation to provide clues to the player about the specific location. If Kahuna faces the boat fan, the buzzing sound will be amplified, and will occur in both the left and right speakers. When Kahuna turns to the left to explore the engine room, the mechanical sound will trail off and pan to the right speaker only, signifying to the player that the fan is at a distance, and is to the right of the player within the gaming environment. These aural cues, establishing spatial and binaural localization, serve to familiarize and orient the player within the surrogate environment, and, combined with visual components, can aid the player in the quest to complete the required objectives to advance the story.

### **5.2.3 SOCOM: Interactive Audio**

SOCOM illustrates how dynamic audio shifts in real-time can enhance the interactive, participatory nature of gameplay. Through interactive considerations in speech, music and sound effects, the audio in SOCOM functions to create a vivid, virtual environment, and also provides ongoing feedback to the player.

Speech components in SOCOM are culturally specific with timbral variations in U.S. ally characters, and with enemy characters speaking in their native tongue, made possible through realistic accents employed by the voice actors. To ensure the variety of enemy voices, there are five different voices assigned to the adversaries to avoid repetition, which is a technique that differs from most FPS games that commonly feature the same voice and identical appearance for most rival characters. Because communication is integral for Navy SEALs, issuing voice commands allows the player to command teammates and have them respond to add to the

realism, and the participatory nature of this interactive auditory experience. The speech component of SOCOM is its most unique, interactive audio feature, in that the peripheral "USB headset" corresponds to a "speech recognition system" within the game, in order to recognize verbal commands spoken by the player.<sup>227</sup> This feature opens new doors to interactivity by allowing the player to issue tactical voice commands to get team members to accomplish specific actions in the gaming environment. The speech recognition system employed by the game acknowledges key phrases spoken by the user, and the other SEAL soldiers will adapt their strategies accordingly. This element serves to immerse the player in a truly interactive entertainment experience, in that the user is dictating the action. The game will recognize such directional phrases as "Hold position," "Follow," "Cover area," and "Attack to crosshairs," as well as attack commands like "Ambush," "Fire at will," "Hold fire," "Deploy grenade," etc. The speech recognition system is not without its flaws, sometimes misinterpreting words that are not pronounced clearly, but is quite accurate, overall, in decoding linguistic variables. The communicative interactivity is further validated when the virtual characters respond. The feedback principle, as explained in chapter one, is a functional element which makes interactivity between humans and machines possible. When the player instructs the Bravo unit: "Bravo, fire at will," Jester will reply "Yes sir, I'm going Maverick" acknowledging that the command has been received; when the player says "Able, follow," Boomer will answer "Right behind you, sir," thereby confirming the demand. In addition to this signal-response feedback relationship, the SEAL team members will periodically offer hints or warnings to cue or coax the player to take a specific action to forward the gameplay after an extended period. For example, in the event of an enemy approaching off screen, Boomer may caution the player "I think they suspect something, sir," while a prolonged delay in accomplishing a task may incite Spectre to blatantly ask "Do you think we should devise a better plan of attack, sir?" or, more subtly, "Waiting on your command." The communication is ongoing, and has been designed to adapt to the player's style of play, speaking more frequently if the player does not act quickly enough, or staying silent if the player seems familiar with how to master the level. Realism is further enhanced by randomized speech, that stores in memory the phrases that have been spoken so that character speech is repeated less often. In online play, human-machine interactivity is replaced by interacting first-hand with other human players to coordinate tactics and strategies to accomplish collaborative missions.

The music ranges from cacophonous, frantic bursts, to gentle, relaxing melodies to signify shifts in atmosphere, and stays consistent with the level of urgency the player must adopt

---

<sup>227</sup> Sony Computer Entertainment Inc. "Behind the Scenes, SOCOM: U.S. Navy Seals": *PlayStation® Underground Jampack*. © 2002 SCEAI.

in a particular gamestate. Aaron Marks likens this technique of suspenseful video game music to film scoring, which is implemented through interactive scoring to enhance the emotional impact for the intended mood of the scene and/or gamestate. Marks comments on the nature of scoring techniques to serve adaptive functions:

Fade-in and fade-out music cues work very well and are used liberally to make adaptive audio applications succeed. Music that begins with a sharp attack is also beneficial – especially when the game mood swings quickly from peaceful exploration to combat mode or when a bad guy jumps out of the bushes... catching [the player] completely off guard.<sup>228</sup>

For example, in encountering a new level the novice player may generally explore and become familiar with the environment before proceeding with caution. The music will adapt to the player's style, and will maintain low amplitude and a suspenseful melody to signify the stealthy approach. If the player is spotted, the rhythm and melody will intensify slightly to signal a shift in gamestate, prompting a greater sense of urgency and for the player to be on guard. Once the player is fired upon, a sudden dynamic shift in the music will occur to announce a game event, and the score will feature a chaotic, frantic, loud and fast-paced segment, signaling alarm to the player and instigating action. The high intensity and energetic nature of the music signals panic, and startles the player. When the player is the one to initiate the attack, the music will intensify more gradually to adapt to the higher intensity of gameplay, but does not serve to startle the player, but rather evolve with the nature of the gamestate. Once the shooting subsides, the music dies down accordingly, crossfading into the initial suspenseful, low-key melody to signify that the player is no longer in a danger zone and can relax. The interactive score fluctuates to adapt the appropriate musical mood to the style of play and corresponds to physical aspects within virtual environment. The intended experience becomes that of participating as the protagonist in an interactive, suspenseful, action-war movie. The emotionally-charged, dynamic score evolves with gameplay to intensify the mood, and takes on a life of its own.

The sound effects, although they are the least interactive of the auditory systems in the game, act as atmospheric elements to complement the adaptive musical score. The sound effects in SOCOM are sample-based by nature, and occur in direct response with changes in the environment rather than adapting to gameplay. This is not to say they do not play an essential role in the interactive audio construct. Sound effects, such as footsteps, will vary according to the user's style of play. For example, footsteps will either increase or decrease, correspondingly in amplitude and frequency, according to the velocity of the character. Sounds are also altered to correspond with a given acoustic space in the virtual environment, making them context-

sensitive. For example, a gunshot's sound envelope will trail off gradually in an outdoor setting, with a reverberant effect added to prolong the decay of the sound to be consistent with the acoustic design. The same gunshot will have a sharper attack and quick decay with little or no reverb in an interior environment to symbolize the absorptive effect that would occur as a result of the surrounding, virtual walls confining the acoustic space. Subtle sound effects can play a vital role in increasing dramatic tension as well. An example of this can be heard when the player uses the scope to zoom in for a sniper attack on an unsuspecting enemy. When the extreme zoom is viewed on the screen, the audio responds with a heartbeat sound to simulate the thumping in one's ears and enhancing the tension one might feel when having to make a critical, precision shot. The visual and aural components are complemented by a tactile sensation of the keypad rumbling in the player's hand. All of these sensations perceived simultaneously by the player are designed to augment the sense of urgency, tension and the need for precision intended by designers.

The end result is intended to be an intensely engaging, militant atmosphere that provides the player with the sensation of taking part in a riveting war movie like *Saving Private Ryan*.<sup>229</sup> While I have deconstructed and segregated the elements of speech, music and sound effects to determine their functional and interactive significance in this analysis, the elements function well as a whole in the game's overall audio structure. The outcome is designed to create a highly immersive interactive soundtrack that attempts to plunge the user into the fantasy world, for the purpose of creating an intense, realistic military team combat experience with nearly limitless replay value.

---

<sup>228</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.192

<sup>229</sup> McCarthy, Ryan. "SOCOM: U.S. Navy Seals." PS2 Staff Review, 2002.

<http://www.psx2central.com/reviews/psx2socom.htm>.



### 5.3 Game 2: SSX Tricky

*SSX Tricky*, an extreme snowboarding game, is a prime example of the Fantasy Sports genre. The goal in the game is to have fun, improving your “tricky” style throughout gameplay by completing multiple variations of advanced *Uber tricks* that defy real-world principles, and take extreme snowboarding to a fantasy level of the surreal. Uber tricks are defined in the game instruction manual as “an all out trickfest,” taking airborne stunts “to the next level with high-flyin’, gravity-defying “Uber” trick(s) that will leave your competitors in awe.”<sup>230</sup> The game is a sequel to the original *SSX* snowboarding game released a year prior, which features many of the same levels and characters, but also incorporates more interactive audio capabilities, superior sound and graphics, and celebrity voice actors. As will be seen in comparison to *MVP Baseball 2003* in the fourth section of this chapter, the audio functions in a very different manner to enhance the fun of an extreme sports game, intending to be “over the top” with exaggerated graphics and sound. Conversely, as elucidated in chapter three, Sports Simulation games features audio and visuals that are necessarily toned down in order to replicate the real-life sporting experience. *SSX Tricky* epitomizes the newest developments for implementing interactive audio in video games, especially with respect to interactive music. The game makes use of DTS (Digital Theatre Systems Inc.) Interactive and SurroundSoundScape audio technologies, varying the music in real-time to correspond with game events, and is only the second game in history to support in-game digital surround sound capabilities (the first being *NHL 2002*).<sup>231</sup> DTS technology makes digitally superior movie-theatre sound clarity, depth, dynamic range and high resolution possible for home console gaming, by internally compressing digital surround sound data into manageable sizes for home console audio playback.<sup>232</sup> To appreciate the surround sound experience of *SSX Tricky* one would need four speakers (for quadrasonic transduction), a DTS-capable receiver, an optical cable to connect the PS2 to the digital receiver, and an optional subwoofer to make use of the game's low frequency effect (LFE).<sup>233</sup> The game audio can also be set to mono or stereo sound for more limited receivers. As can be seen in the *SSX Tricky* sound count sheet (Appendix C), the game incorporates a variety of sound effects, music samples, and voices within the game. What cannot be simply catalogued are the complex virtual sonic environments, and the interactive, real-time function of these sounds within the soundscapes, which will be explored in sections 4.3.2 and 4.3.4.

---

<sup>230</sup> ©2001 Electronic Arts, Inc. *SSX Tricky* Video Game Instruction Manual, p.17

<sup>231</sup> O'Neill, Cliff. "SSX Tricky Game Review." Sports Gaming Network, 2001.  
[http://www.sgn.cc/other/ssx\\_tricky\\_ps2/review\\_ps2.shtml](http://www.sgn.cc/other/ssx_tricky_ps2/review_ps2.shtml).

<sup>232</sup> DENON Ltd., 2003. FAQ Website: *DTS* <http://denon.jp/asia/faq/2-8.html>.

### 5.3.1 SSX Tricky: Speech, Music & SFX

The fast-paced audio environment created in *SSX Tricky* adrenalizes and excites the player with energetic, new music spanning a variety of genres, high-energy announcers and celebrity voices, and realistic, variable sound effects; making the game truly immersive and exuberant. Before examining these sounds within the sonic environment and their interactive function, I will first highlight some of the key acoustic, spoken and musical constituents.

The speech components of *SSX Tricky* are in the form of announcer and character voices, as catalogued in the sound count sheet (Appendix C). The voices are acoustically distinctive from one another, each having unique pitch ranges and timbral characteristics, which allows for a hi-fi soundscape when voice samples overlap during gameplay. For example, the voice of Seeiah will feature a squeaky, high-pitched, raspy tone colouring, while the voice of Luther will have husky, guttural characteristics, all contrasting with the MC voice, comprised of low frequencies and smooth timbre. Linguistics hold less emphasis than paralinguistic components of speech in *SSX Tricky*, as the experience is meant to be fun and exhilarating rather than an instructional lesson or learning experience. The players seek to improve their style of play in order to progress to expert levels and unlock special features. The witty linguistic components suit the hip style of the game, incorporating phrases that will resonate with teens and adolescents. Buzz words like "sweet", "dope" and "funky-fresh" will be spoken to encourage the player, while negative expressions such as "lame", "brutal" or the blatant "this sucks" will signal negativity to the player. Linguistically, all of the one-liner speech phrases are appropriate to the personalities of the characters speaking them, and reflect the style of play. The paralinguistic connotations of speech dominate the semantic meaning of the phrases. The character of Eddie shouting a triumphant "Woohoo!" when landing a complicated trick, or a dismayed "Arg!" when falling down, has a greater impact for the player than a verbal explanation as to the effectiveness of a given maneuver. Also the word "great" can have both a positive or negative connotation depending on how it is spoken by a character in the game, either implying praise or sarcasm. If a player says "come on," it either signals outright frustration, or urges that the player needs to speed up depending on the intonation. The word "bam" spoken by the MC has no linguistic meaning in and of itself, but said in a high-amplitude, forceful and energetic manner by MC Rahzel implies a sense of accomplishment to the player upon discovering a new trick. The MC voice was chosen because of his "beat-boxing" capabilities, which is a technique of using the voice as a percussive

---

<sup>233</sup> O'Neill, Cliff. "SSX Tricky Game Review." Sports Gaming Network, 2001. [http://www.sgn.cc/other/ssx\\_tricky\\_ps2/review\\_ps2.shtml](http://www.sgn.cc/other/ssx_tricky_ps2/review_ps2.shtml).

instrument much like a drum machine. The voice of Rahzel (self-proclaimed as a human sound effects machine) is a unique element that combines all three of the systems within the continuum, serving as a speech element with the words spoken, a musical element with sung notes and percussive noises, and sound effects with his various sounds, noises and utterances. Interestingly the MC voice ties the speech, music and sound effects systems together in one, in that there are times when Rahzel's voice changes from words to musical sounds to abstract noises almost simultaneously. I would suggest that Rahzel's voice serves as a metaphor for what the sound designers at EA were attempting with this innovative approach to character voices and speech in SSX Tricky: it combines, blends and juxtaposes the acoustic elements; morphing them together to create a seamless, lively soundscape that ebbs and flows in real-time compliance with the player's actions.

The musical score for SSX Tricky incorporates many styles of music including funk, disco, rock, heavy metal, techno, electronica, rap and hip-hop, featuring songs licensed by big-name, contemporary artists, and all contributing to the fast pacing of the game. The SSX Tricky sound count sheet (Appendix C) contains a list of the songs contained in the game. The varied soundtrack consists of energetic songs, consisting of those licensed by popular artists and original compositions by EA composer and designer John Morgan. A number of musical samples also insert themselves throughout the game to emphasize "tricky" maneuvers. Each character has its own theme song that will sound while customizing player attributes, or when the NIS "victory" sequence plays to reward the player for acquiring an advanced player ranking. Each venue offers the player a few songs to select from to suit the user's preference, so as not to irritate the player. This serves to enhance the replay value of each level, and varies the soundtrack when the player needs to continually replay a given level. As the SSX Tricky sound count sheet (Appendix C) indicates, there are 37 different songs that have been assigned to specific venues and characters in the game so that songs do not loop the same section over and over again as in other games, thereby avoiding monotony. This is accomplished by having different sections of songs play in alternating orders to maintain a varied soundtrack that will play differently each time. Although the songs remain the same for each level and repetition is unavoidable (as there are only a select number of musical combinations that can exist before a sequence is duplicated), redundancy is kept to a minimum by having them occur at different intervals in the cycle. Much like a jukebox or CD player set on random play, the result is a soundtrack that is designed to keep the player/listener on edge, never knowing what to expect from the score, which mirrors the unpredictable nature of SSX Tricky gameplay. As will be discussed later in this analysis, the

unpredictable nature of a real-time soundtrack that adapts itself to random gamestates serves as an inventive way to make interactive music that reflects the user's style of play in SSX Tricky.

SSX Tricky features a wide array of vivid and dynamic sound effects to submerge the player into a larger-than-life, extreme gaming experience. High-quality ambient sound samples are incorporated, covering animal (birds, crows, owls, wolves), natural (wind, water, snow), human (crowd) and mechanized sounds (snow machine, snowmobile, chairlift). The primary sounds heard are those of the first-person character's snowboard gliding and carving the trails, maneuvering through the game's numerous, challenging tracks and courses. The board sounds, as listed in the sound count sheet (Appendix C), contain multiple sounds that overlap and layer in real-time according to the physical context and gamestate. The varied "whooshing" sounds of carving, grinding and jumping are vivid, crisp and serve to stimulate the player's auditory system. Sound effects correspond with visuals to simulate realistic sounds analogous to the environment. For example, the harsh timbre of the snowboard scraping on dirt or rock will be represented with different sonic characteristics from the contrasting smooth tones of the board gliding through fresh snow. Sound effects are often exaggerated to augment the impression it will have on the player, such as the dramatic, high amplitude, modulated "swash" sound that is heard when jumping, elevating the character high above the track below. The sound of the swirling wind also strives for high impact with exaggerated, digitally-processed wisps that effectively embellish the impression that the player's character is soaring at high altitudes. Crowd sounds are acutely composed of several layers of ambient chatter that would be heard at a sporting event. Much as the courses are exaggerated, the crowd size is also inflated, visually appearing as if tens of thousands of spectators have gathered to witness the events. The audio mimics this by adding chorus and reverb effects to make the crowd sound much larger, and seemingly widespread. Very few (if any) individual voices stand out in the crowd cheering, making it sound much like a wall of broadband noise, varying in intensity and making use of band-pass filtering through the frequency spectrum to create an effect much like the oscillations of ocean waves. The loud, distinct, grating sounds of a collision with an object are well-defined and noticeable, taking precedence over other sounds to serve as a jarring and abrupt end to any progress or momentum instigated by the player's actions. Sounds of glass shattering and fireworks squealing and exploding contribute nicely to the effect of large crowds cheering to enhance the event being conveyed in the game. These sounds incorporate high frequencies and bright tone colouring to stand out in the mix to signal a game event, and are sure to be heard over other sounds in the overall audio mix. Together with the dynamic voices and music, the game audio serves as a

dynamically enveloping component contributing suitably to the “in-your-face” nature of the game.

### 5.3.2 SSX Tricky: The Virtual Sonic Environment

The auditory experience created for SSX Tricky sets a new standard for video game audio. The DTS audio construct allows for full immersion in the game, as music, voices and sound effects are dispersed through different channels in real-time. The virtual sonic environment is complex and vivid, encompassing the player within the surround audio environment. The quadraphonic effect creates the illusion that rival snowboarders are approaching from behind the player, while crowds cheer from various angles, and ambient, background sound effects occur in relation to the player's listener position within the game's soundscape. It is interesting to note that in 2-player mode, the audio attempts to position both listeners within the quadraphonic environment, but with specific audio signals pertinent to their character's gamestate and position. The music will remain the same for both players, although their specific sounds and voices appropriate to the player's character will be directionally specific with respect to the stereophonic or quadraphonic channels. The 1<sup>st</sup> player, whose character is contained in the left half of the screen, will have the corresponding sound effects and speech appropriate to the character and its positioning within the environment transmitted from the left speakers to preserve a balanced audio/visual relationship. The 2<sup>nd</sup> player, on the contrary, will have character-specific sounds activate in the right channels to have the desired panning effect for the user. The audio environment creates a directional effect that can immerse the player to create a spontaneous real-time experience. As Cliff O'Neill writes in his review of SSX Tricky for *Sports Gaming Network*, "[SSX] Tricky has more directional effects and stronger bass than many action-packed DVD movies encoded in Dolby Digital 5.1!"<sup>234</sup>

Because the character voices and MC voices are spoken by well-known celebrities, they function as recognizable elements within the soundscape that players can identify with. The other announcer voices summarize or set up the action between races, while the character voices speak during racing gameplay, character selection, and make comments following the races to signal satisfaction or disappointment to the player. The function of voices and speech is to guide the player through linguistic and paralinguistic cues. Just as the sound effects work to provide auditory hints about the environment, the speech components of SSX Tricky serve to provide hints and tips to the player, ranging from subtle vocal allusions to blatant verbal commands. The character voices serve as both linguistic and paralinguistic indicators, becoming an internal

dialogue for the player to either provide encouragement or criticism, and to signal celebration or disappointment. Through positive reinforcement phrases like Eddie's "Right on, baby", Seeiah's "Check me out, I'm superfly", or Mac's "Unstoppable! Keepin' it real, yo"; the player dialogue functions to motivate the player. On the other hand, the voices can offer constructive criticism such as Mac's hint "You gotta focus!"; can signal encouragement with phrases like Seeiah's "Fall down, get up, get over it!"; or mirror the user's outright frustration with expressions such as Eddie's "Wake me up when I'm winning", or "Yup, this sucks, alright." The celebrity voices are high-energy performances that serve to excite and engage the player as well as providing comic relief from the intensity of gameplay, and are meant to be likeable and are designed to evoke a sense of companionship for the player.

The MC/announcer voice of "Rahzel" acts as an ongoing play-by-play of the racing action, guiding players with positive feedback or negative taunts and jeers. The MC will serve to egg the player on or encourage the player to adapt the style of play. Rahzel has a booming male voice made up of mostly low frequencies that dominates the underlying character voices. Because Rahzel is well-known as a DJ and beat-boxing champion in hip hop rap music, his voice acts as a soundmark that the player listens for to provide recognition and cues to guide the gameplay. In addition, Rahzel's status as a live DJ in the real world enhances the realism and credibility of his voice as the "Voice of God" that the players must abide by in order to excel in the game and elevate their style of play to become a competitive boarder. The MC voice is intended to guide the player with constant verbal feedback that takes on a variety of roles: to motivate, encourage and reward the player; to offer hints, constructive criticism, coaching and tips to improve the player's style; to signal and to act as a real-time deejay or play-by-play announcer to comment on the player's actions as they occur. The linguistic variables must be numerous and specifically integrated to be effective, in that the MC voice must function as announcer, coach, assistant, friend and DJ. Updates on the player's position must feel sporadic, and compliments and criticism of the player's progress and overall performance must be realistic. For example, if a player lands a complicated trick, the MC will remark directly to the player "Oh yeah, that's what the crowd wants!", "You're the show-stopper", or "Call your momma in the room and show her how great you are." In attempts to motivate players, but at the same time challenge them to try more difficult moves, Rahzel will say "Now you're getting it, let's see your signature move." When a player fails to vary their style, or repeats maneuvers during gameplay, the MC voice will criticize the player, making sure not to be insulting with phrases like "We need to see a new trick", "Again?" or "That trick is like a bad habit." Although repetition does occur because of the

---

<sup>234</sup> Ibid.

constant bombardment of one-liners, the natural flow of speech is generally preserved, and excessive reiteration is avoided by the inclusion of paralinguistic utterances such as groans, grunts, hoots and hollering by the MC in between phrases.

The musical soundtrack consists of 37 songs that can be appreciated separately as pieces of music (that can be played solely for listening purposes in the “Jukebox” section of the options menu), but arrange themselves together to make up the gaming environment to instigate a mood, keeping pace with the game and player style. Because the pop artists and record companies lending their songs to the game prefer that their original mix not be tampered with in the game, they are generally incorporated in their original form, and sampled within the game during non-interactive sequences. The non-adaptive songs still serve an important function in the gaming environment, in maintaining the energetic mood of the game, as well as tying the game music to real-world pop music culture. The licensed songs by popular artists are recognizable to the player demographic, and serve as musical soundmarks in that they are ties to the real-world environment. The pop songs are used either as individual character themes, or to reward the player in the game by playing the chorus during the celebration NIS when a milestone in the game is reached, and a new player ranking is earned. The adaptable songs (that play during racing gameplay), layering musical streams in real-time with the player’s style, are the compositions of EA composer John Morgan. Having both designed the music and implemented his songs into the game, the interactive potential of a self-building composition that adapts itself to real-time gameplay comes through in the soundtrack. Music will range in style to vary the pacing of particular levels, and will shift according to gamestates. Different songs will play for specific players during non-interactive sequences (typically songs by pop artists), and select songs will be assigned to specific venues and levels (those of in-house composer and designer John Morgan). The music functions to provide a mood, but also to cue the player to take specific actions, as is embodied primarily in the “It’s Tricky” sample taken from the RUN DMC song of the same name, originally released in the rap group’s 1986 album, *Raising Hell*.<sup>235</sup> The energetic song reflects the pacing of the game, and is identified as its recurring theme, giving it a retro feel. The “It’s Tricky” sample serves multiple functions, the most impressive being its interactive role, which will be discussed shortly, but also as reward music, and as a signifier. A key role that the sample plays in the SSX Tricky soundscape is that it becomes a soundmark for the player, in that it identifies and signifies a heightened gamestate each time the song is heard. Not only is the song identifiable as the popular 80’s hit theme song, but it signals to the player that the Tricky Meter has reached its maximum boost level, and that the player only has twenty seconds to complete an

uber trick. A longer sample of the “It’s Tricky” chorus will sound each time a standard trick is performed to maintain the Tricky Meter; reminding, anticipating and coaxing the player to complete an uber trick. The sample will play along with the siren sound to reward the player with a frenzied cacophony of noises much like the auditory feedback that would result in winning a prize at a fair. However, the sample serves yet another function with a negative connotation in the form of a low frequency, descending version of the same soundmark. If the player fails to perform the uber trick in the allotted twenty-second time limit, the pitch of the sample will lower a few octaves to simulate a DJ reducing the speed of a turntable. The implied meaning of the descending pitch is to symbolize a loss of energy as well as a disappointment. The “It’s Tricky” soundmark then, embodying identification and familiarity for the player, is a multi-purpose sample that is designed to provide a sense of reward, defeat, or heightened excitement and anticipation for the player, which is fluently beat-synchronized to rhythmically coincide with the tempo established by the current background soundtrack of the level. The fact that the soundmark mixes into the background music whenever the Tricky Meter is powered up allows for a seamless cue to action that creates a remix each time it is inserted into the score, and does not clash or disrupt the flow of the soundtrack; thereby allowing for the fluid transition through evolving gamestates without any jarring interruption.

The function of sound effects in SSX Tricky is to enhance the fun of the play experience with exaggerated, over-the-top sounds to either emphasize keynotes in the environment, or to highlight the tricks being performed. Variable sounds of the player’s snowboard traversing through the course are constantly shifting in timbre from coarse sounds of grinding on slick ice compared with the soft “whoosh” of carving through virgin snow. As environmental indicators, mechanical sounds of snowmachines will signal a warning to the player that the user is in close proximity to an object that could serve as an obstruction. Each venue in the game has distinctive environmental sound effects ranging from the soothing, ambient sounds of natural waterfalls in the Aloha Ice Jam level to the chaotic, electrical sounds of the pinball environment of the Tokyo Megaplex level. Environmental and animal sounds specific to each stage add to the realism of the intended atmosphere, giving each track its own, exclusive, auditory signature, with different bird and animal samples sounding intermittently to signify a different setting from the previous stage. The natural and animals sounds are seemingly unaltered by effects, leaving these high quality sounds to sound as they would in a realistic environment, which creates a nice contrast with the DSP manipulations applied with many of the ‘action’ sounds.

---

<sup>235</sup> Ibid.



Bells, sirens and jingles are pronounced tones incorporating multiple high-frequencies to signal or alert the player or cue a specific action. A high-frequency drone burst will sound when the player reaches the course boundary or becomes stuck and unable to move. The sound functions to represent a time lapse, as there is a momentary pause in gameplay while the character is repositioned and re-inserted onto the course in the same position as before the mishap or accidental deviation occurred. DSP manipulation works to overstate sounds and make them larger and more impactful than in the real world soundscape. An example of this presents itself when the player times a jump perfectly at a high speed and achieves “maximum hang time,” allowing the player to perform intricate stunts in the air for bonus points while soaring for extended periods before hitting the ground. To enhance the ambient whistling of swirling winds to simulate high altitudes, a DSP inflection is added in real-time by audio algorithms in the game to add a flange effect and EQ filtering to add a warbling, phasing effect to the wind sound, making it sound exaggerated and fantastic. The effect is designed to be subtle, but functions as a recurring keynote sound for the purposes of signaling to the player that a high enough altitude has been reached to attempt a more intricate and complicated maneuver for a high score before landing and resuming the race.

### **5.3.3 SSX Tricky: Interactive Audio**

Audio interactivity in SSX Tricky is a technological achievement that uses detailed, real-time algorithms that treat sound in inventive and innovative ways. Although speech, music and sound effects are elements that all function as one interactive unit, designed to make for an interactive audio experience, the components can be segregated to pinpoint their relationship to the player in the overall audio construct.

The speech components in SSX Tricky were revamped from its predecessor, *SSX*, which was released only a year earlier. The original version had less emphasis on vocal performance (which was improved with the addition of celebrity actor voices), and featured venue-specific pre-race announcers rather than an anchor who narrates details of each course before the beginning of each race. Not only are there numerous phrases and expressions spoken by each character in the game, but the intonations and paralinguistic emphasis of key words that merit repeating have multiple variations to uphold adaptivity within different gamestates. The paralinguistic intonation employed may also change to simulate how the announcer might react to the style of play, speaking the phrase in a given scenario. For example, Rahzel will say “Dope!” numerous times to signal completion of a standard trick. However, the second time it is repeated, the pitch may be altered slightly to simulate how the MC might vary his spoken intonation if he

were to repeat the same word. Linguistically, the speech A.I. in the game applies appropriate words, or can even say the same thing in a variety of ways. For instance, if a player repeats the same trick twice in a row, the MC voice will tell the player “The crowd wants to see a new trick.” If the move is repeated again, another MC sample will warn “The same trick only equals half the points.” If the trend of the player repeating continues, the comments will escalate to scolding, suggesting that the announcer is becoming bored and frustrated, with remarks like “Too much repetition”, “That trick is like a bad habit”, “Again?” or even the stern warning “Don’t make me say it again.” Consequently, if the player lands a different move successfully, the MC will acknowledge that the player tried something new, congratulating the player with phrases like “Now you’re getting it”, “That’s what the crowd wants” or “Now you’re riding with style.” While the phrases being spoken are still being chosen by the audio algorithms from a sound bank of speech samples, the interactive component is that the audio A.I. is able to prioritize, and seems to know which samples to insert at the appropriate time in order to respond suitably to any possibility.

In addition to applying the appropriate linguistic samples and paralinguistic intonation, certain speech samples are modulated on the fly to symbolize a heightened gamestate. For example, when the player does enough tricks to fill the Tricky Meter to full boost, the word “Dope” spoken by the announcer may be adjusted slightly to incorporate a reverb and flange DSP effect to enhance it. There is also an instance where the word is interrupted by a turntable scratching sound, as if it were being played by a live DJ at the venue, making the end result “Dope-D-D-D-Do-Dope!” The MC voice becomes modified with scratching sounds as well as applying reverb, EQ parameters and filtering to make the words sound synthesized and almost robotic, signifying that the player’s style has evolved to a new echelon.

In addition to the linguistic and paralinguistic components of speech to convey meaning to the player, the spatial, digital surround soundscape introduces another dimension to the interactive perception of the game’s audio construct. Localization within the environment and in relation to other characters becomes further enhanced when the quadrasonic environment is taken into account. In addition to directional, environmental sound effects, such as spectators cheering from multiple directions in relation to the player’s position on the course, the speech components during gameplay interact with the player within the environmental setting. For the first time in a snowboarding game, the player can actually hear opponents approaching from the left, right, front, or even directly behind the player without first seeing the rival characters on the

screen.<sup>236</sup> This adds a new element of interactivity for the player to engage with the speech components of the game, as the characters can be heard shouting comments, taunts, threats or celebrations from the rear speakers when they are approaching from behind. The effect is one that encompasses players by surrounding them in a wall of sound, making the sonic environment multi-dimensional and vivid, rather than being isolated in two stereophonic speakers contained in a television monitor, located directly in front of the user.

Interactive music, with regard to *SSX Tricky*, functions as the most significant, innovative and immersing component of the audio environment. The musical score is an adaptive soundtrack that builds itself through instrumental layering in relation to the player's style of play. The music will reflect the pacing of the race, mix layers into the score in real-time as the player's performance improves, and will even fade down in amplitude during a jump to signify anticipation, and to represent that the character is soaring above the speakers in the gaming environment. The soundtrack for each level will begin with a basic rhythm and melody, and will build itself in real-time in response to the success of the player, and in relation to the level of the Tricky Meter. Although this technique of using multiple music streams at the same time in order for instruments to fall in and out of the soundtrack depending on how well the player is doing has been employed sparingly in the past (e.g. *Parappa the Rapper*), the technique has never been applied to a fantasy sports game, and certainly never in such an interactive way that responds directly to a player's actions. While some of the more popular songs included in the game are inserted to enhance the pop culture feel of the game, and are by nature non-interactive, the original compositions of John Morgan that make up the core of the interactive soundtrack are made to engage with the player and adapt to user input. Interactive music algorithms in the game will build or take away instrumental layers of the soundtrack to convey positive or negative developments in gameplay. The entire soundtrack, in this case, becomes reward music if the player succeeds in accomplishing a variety of tricks. In addition to accumulating a high score, the player is rewarded musically with a tumultuous opus that the user has, in a sense, created by completing multiple maneuvers in the game. The sense of accomplishment is reflected in the complex symphony of harmonious melodies that are constructed. Let me explain how this works in the game.

In the second stage of the game entitled "Snowdream," the level begins with a basic score consisting of only a bassline and base drum track. As the player accomplishes tricks, accumulating notches on the Tricky Meter, the soundtrack will continue to increase in variety and

---

<sup>236</sup> O'Neill, Cliff. "SSX Tricky Game Review." Sports Gaming Network, 2001.  
[http://www.sgn.cc/other/ssx\\_tricky\\_ps2/review\\_ps2.shtml](http://www.sgn.cc/other/ssx_tricky_ps2/review_ps2.shtml).

complexity, first adding a rhythm guitar and a more varied percussion loop, creating an augmentation in the intensity of the gamestate. The more the player varies the style of play in Snowdream, and the more rungs that are added to the Tricky Meter to boost gameplay, the more layers are added to the soundtrack, successively appending bassline variations, drum-fills, vocal samples, record-scratching, synthesizer overlays and flute and organ segments. Once the Tricky Meter reaches the highest level, a bell sound effect will ring twice, coupled with a siren sound, and the “It’s Tricky” soundmark will ring in to cue the player that the 20-second countdown has begun to perform uber tricks. At this stage the music will be at its highest intensity, incorporating the maximum number of music streams, with multiple layers of instruments all beat-synched together to create a cacophonous and chaotic piece of music in time with the pacing of the gamestate, signifying the urgency for the player to use the allotted time to perform advanced maneuvers to collect bonus points. Each time the player accomplishes an uber trick, the siren and theme-song soundmark sequence will repeat to symbolize that the 20-second timer has reset, and serves to encourage the player to try another sequence of tricky maneuvers and uber tricks. After the first uber trick is accomplished, the atmosphere becomes even more adrenalized, which is represented through DSP modulation of the musical score. To further intensify the soundtrack, the music will digitally morph itself into a warbling, effects-processed version of the final mix to signify the summit of the heightened gamestate. This final version of the song will continue until the Trick Meter level decreases once the player crashes, or fails to complete a trick within the 20-second time limit. When the player’s performance dwindles, the music streams will systematically remove themselves from the mix, leaving a more basic soundtrack that diminishes the overall intensity of gameplay, urging the user to aim higher to recreate the intensity level that was previously created by accomplishing an assortment of tricks.

Levels of accomplishment or disappointment are designed to be communicated to the player through nuances and gestures in the interactive musical score in SSX Tricky. The real-time mixing system will insert or subtract musical layers, which reaches a climax with the “It’s Tricky” theme as the final addition to the resulting ear candy. SSX Tricky may set the bar higher for interactive audio in fantasy sports games, and may very well change the way video games in general are heard, perhaps one day making standard stereo game sound a thing of the past.

Sound effects will overlap and juxtapose with one another in real-time to correspond with the player’s actions. Sound effects such as the bell sound indicate that the player is progressing, and has earned another notch on the “tricky meter,” bringing the user closer to tricky status. A high-pitched siren will sound, coupled with a low frequency “tricky” uttered by the MC/announcer, and the bell will ring twice to signal that the player has reached the highest level

on the tricky meter, and can begin to attempt uber tricks for high scoring. These sounds are integral to signal precise timing for the player, as there may only be a split second for the player to react and accomplish a trick when traveling at high speeds and approaching a jump. The sounds must occur without delay to signal urgency to the player. The audio is designed to cue the player to action, while the player's eyes remain fixed on the first-person character, so as not to crash or deviate from the course. Natural sounds of birds chirping or wolves howling, in addition to the gradual fade out of crowd ambience, signals that the player has veered off course. The crowd sounds will correspond to spatial location within the environment, and will get louder and have directional panning adjustments in relation to the player's distance from the crowd. Obstacles such as tree branches will make a "thwap" sound if struck, or a more gentle brushing sound if grazed, depending on the player's proximity to them.

The sounds will also vary depending on variables such as a player's size and weight, as well as the velocity of the player hurtling down the track. For example, the character of Kaori takes the form of a small Japanese girl, standing five feet tall and weighing less than a hundred pounds, while the character of Luther is an overweight American male, measuring six feet, five inches and weighing closer to three-hundred pounds.<sup>237</sup> When Kaori speeds down the trails, the sound of her board against the frozen ground is higher in pitch to correspond with her size, while Luther's board will have a deeper, coarse tone to represent his weight on the board. When the character crashes, into a tree for example, the "thud" sound of the impact will be lower in amplitude and higher in pitch for the character of Kaori than that of Luther, and will also fluctuate depending on speed variables. Sound effects will overlap and juxtapose in real-time, but will prioritize so that sounds will decrease or increase in amplitude related to other competing sounds. In the event of fireworks exploding, the firework sounds will dominate as the player's character passes them, but their amplitude will diminish and give way to continuous sounds (such as the board carving sound) once the fireworks are in the distance. The result is an ambient soundscape that changes dynamically and spatially depending on the player's position within the environment and style of play in particular gamestates. "Action" SFX will also adapt to gameplay by layering different versions of sounds together to achieve the desired effect depending on the gamestate. For example, if the character jumps off a high cliff the user will have ample time to perform a wide array of tricks while in mid-air. Each time a trick is performed a "buzz" sound will activate to signal the completion of one move from the trick book, which will sound each time an additional trick is performed to acknowledge the combination of simultaneous moves to accumulate bonus points.

## 5.4 Game 3: MVP Baseball 2003

*MVP Baseball 2003* is an excellent example of the modern sports simulation, or *sim* game, in which elements of speech, music and SFX are used to simulate a realistic and engaging sporting experience for the player. The aim of the sim game is to balance the perspective of a spectator viewing a live telecast and that of a player in the actual gaming context. The audio, feeding off the visuals, attempts to accomplish both of these tasks by mirroring the graphical components and providing atmospheric sounds appropriate to the environment, while creating a life-like, engaging and interactive sonic experience. The audio positions players within the environmental gaming context, and encompasses them with replicated, ambient stadium sounds as players control members of their selected team. While it is clear from the introductory blurb (Game Description: Appendix D) that the emphasis is still on the visuals, emphasizing the “look and feel” of the game to make it “true-to-life,” the sonic environment is quite advanced to create the realistic “feel” of the game with a vivid stadium soundscape environment and pragmatic speech and music implementation that is designed to immerse the player in the sports world embodied in the game. Let us examine the sounds in greater detail before deconstructing their interactive potential and atmospheric function within the artificial environment.

### 5.4.1 MVP Baseball 2003: Speech, Music & SFX

The speech components within the game are personified primarily through the ongoing commentary of real-world San Francisco Giants announcers Mike Krukow and Duane Kuiper. Because of their real-life affiliation with the game of baseball in announcing games for the Giants, the dynamic commentary authenticates the statements made to support gameplay, and, because of their working relationship, Krukow and Kuiper’s comments play off one another seamlessly for a fluid play-by-play simulation. Their voices are clear and articulate, and are superimposed in the overall audio mix to stand out above the game ambience, intended to guide the player through the game. However, the player may wish to decrease the commentary volume in the options menu to simulate more of an in-game atmosphere on the field rather than that of a telecast, depending on the user’s preference. The commentators provide enthusiastic commentary with articulate linguistic emphasis, and pronounced paralinguistic inflections to get their points across to the player. Both are very convincing, in that the play-by-play is delivered in a blatant, factual way, while the colour commentary is spoken in a more convincing manner, intending to

---

<sup>237</sup> ©2001 Electronic Arts, Inc. *SSX Tricky* Video Game, PS2 “Rider Profile: Kaori; Luther”

appeal to the player much as a friend would while venturing a perspective on an event. Speech narration in baseball is not as fast-paced and constant as other sports such as hockey or basketball, which would make it more appropriate for commentators to deliver more long-winded phrases about game events, player attributes, seasonal statistics, team information, etc. The game replicates this well with samples that preserve the natural evolution of speech, and do not sound stilted when stringing together different remarks. Although the names of specific players voiced by the commentators may occasionally be spoken at a different pitch from the rest of the sentence, the speech is fluid and seamless for the most part. The sophistication in the linguistic variables is considerable, in that the announcers do not interrupt one another, and their comments will coincide with the appropriate action taking place on the screen. For example, when the pitcher throws a strike and the batter swings and misses it, the voice of Kuiper will comment on the pitch down to the specific detail of the location and velocity of the selection, remarking "Swung on and missed the split-finger fastball at the knees." Krukow will then ring in with a conversational and often humorous observation that plays off Kuiper's comment like "Ya, there's something missing with that swing, alright: contact!" The fact that Kuiper has commented on specific details about the pitch, coupled with Krukow's sarcastic and witty comment about the lack of contact simulates a realistic commentary that describes exactly what has transpired, and makes it seem as if the baseball analysts are witnessing the game and calling the plays in real-time.

Other than the commentators, the other speech elements in the game are the PA announcer, umpires and voice direction provided by baseball hall-of-famer, Harold Reynolds in tutorial mode. The tutorial voice narrates specific tips and techniques for improving gameplay when the *MVP Tips* menu item is activated, which is an NIS featuring instructions on how to bat or pitch effectively. The voices of the two different umpires call the balls, strikes and foul balls at the plate, and whether the runner is safe or out in the field. Both the home plate umpire and fielding umpire deliver their decisions with clear linguistic emphasis, ensuring that their judgment is audible with enthusiastic and energetic paralinguistic emphasis. The PA announcer voice will inform the crowd (and functions to cue the player) of a new batter approaching the plate, or of a player substitution during gameplay. The voice is delivered in a low frequency tone that has little paralinguistic variation, stating player names in a factual manner.

The musical soundtrack for *MVP Baseball 2003* is comprised of an impressive list of songs from established and up-and-coming pop rock bands whose music is featured, many for the first time, in the game with EA Sports Trax™. The songs range from alternative rock to new-wave metal, complementing the excitement effect evoked by the flashy, fast-paced visuals on the

screen. In featuring forthcoming, licensed songs by some previously unknown bands, EA has taken the initiative to promote artists before they become popular with mainstream listeners. Thereby the music in the game also serves as a promotional tool for artists and record companies to feature new music in games as a vehicle to introduce songs to their target audience in an interactive entertainment context. The new songs that are showcased from established bands function as a source of recognition for the player, as the songs represent familiarity in the user having heard of the bands, giving the game credibility as being current and “cool.” Other musical elements are featured, although sparingly, during gameplay. The classic keyboard sounds that are associated with baseball stadiums are captured in the form of organ samples, which are inserted into the crowd ambience to recreate the build-ups that occur in professional sporting events. Continuous in-game melodies are, by nature, non-existent in sports venues, although player introductions for the home team players are commonly preceded by a rock song to pump up the player and stir up the crowd. This is recreated in MVP Baseball 2003 in the form of specific songs that will play as each new batter for the home team approaches the plate to bat. The songs played during gameplay are understated, and become ambient sounds to contribute to the overall soundscape rather than becoming dominant elements. The songs and organ ditties are treated with reverb effects, and are filtered and compressed to sound as if the music is being transmitted from the PA system within the gaming environment.

Scrupulous attention to small, sonic details serves to enhance the overall realism of gameplay. The only deviation from the rock ‘n roll feel of the game’s soundtrack, aside from the odd three-second Latin song clip introduced as a player intro, is the techno rock soundtrack accompanying the competitive Homerun Showdown gameplay mode. This mode of gameplay serves as an interesting contrast to simulation gameplay, in that it is short compared to the potentially lengthy experience of playing a nine-inning game of baseball, and is fast-paced, contrasting with the drawn-out nature of baseball gameplay. The showdown requires the player to hit homeruns in direct competition with the computer, or against another human challenger, to reach an accumulated distance total faster than the opposing character pictured on the other half of a split-screen. Musically, the Homerun showdown serves as a refreshing contrast to regular gameplay, with a high-speed, energetic techno background designed to exhilarate the player.

The sound effects are arguably the most striking feature of the game’s audio, with clear, dynamic, digital SFX; the most complicated and vivid being the variable crowd sounds. The dynamic crowd sounds in MVP Baseball 2003 are some of the most detailed, realistic and complex audience sound effects ever captured in a game, featuring multiple layers of ambient crowd chatter, cheering, booing, shouting, clapping, chanting and stomping fans rooting for their



home team. The atmosphere replicates a lifelike crowd of spectators that has varying levels of intensity depending on the gamestate. The crowd ambience will shift along with events in gameplay, and occasional crowd voices will be heard above the ambient chatter, shouting supportive statements to their team like “Hit the ball” or “Come on, go for it!” The voices will also heckle the opposing players or criticize the umpire accordingly. Meanwhile, vendors will hawk their wares, tempting fans to purchase anything from hot dogs to sodas, beer to burgers, snowcones to roasted nuts, popcorn to ice cream to lemonade, etc. The variation in the vendors’ speech is considerable, with a wide range of products, and a multitude of different phrases to sell them. For example, I was able to discern three different samples for the popcorn item alone, varying from “Popcoooooorn” to “Popcorn, come on, who wants popcorn?” and furthermore “Popcorn, you want popcorn? We got it! Get your popcorn!” Because of the seemingly endless variety of samples being cycled throughout the game, the player rarely notices any repetition, and if any does occur the samples will be quite spread out, and new phrases may be heard for the first time even after several games are played. Sounds of clapping, stomping, chanting, whistling and screaming and even the occasional foghorn will emanate from the crowd ambience, all adapting and building in accordance with gamestates to provide the proper intensity in the overall atmosphere. Spatial characteristics are evident in the crowd sounds, with different aspects of the crowd ambience sounding from different angles to encompass the player within the environment. In-game SFX specific to baseball are clear and pronounced, with vivid bat and ball sounds during the batting sequence and glove impact and sliding sounds during fielding. Action sounds, namely the striking “crack” of the bat meeting the ball, or the popping sound of the ball being caught in the glove are clearly palpable, each having a sharp attack, and crisp, bright timbre. The harsh, grinding timbre of a runner sliding into a base are replicated, as are the dull “thud” sounds of a ball hitting the outfield wall. Announcer voices are treated with excessive reverb, while umpire and commentator voices are clear and preserved without effect manipulation to maintain clear audibility. Human voice sounds of crowd voices and vendors, streaming in conjunction with PA music and announcer voices, as well as clear, vivid sound effects are designed to work together for the purposes of creating a well-balanced mix to simulate the atmosphere of a real stadium.

#### **5.4.2 MVP Baseball 2003: The Virtual Sonic Environment**

The slogan for MVP Baseball 2003, spoken by multiple baseball stars in NIS in the introduction to the start-up menu, is “EA Sports: If it’s in the game, it’s in the game.”<sup>238</sup> The goal, as implied by the slogan, is to replicate the atmosphere down to the last detail of being in the

middle of the action at a real, big-league ballpark, which is recreated through flashy visuals, but more than any other element, the realism of the environment is brought out through the soundscape. The overall audio design of the game functions to authenticate the artificial environment, which seems to be the underlying purpose of the audioscape, perhaps even more than interactivity.

Although the sports sim gameplay can exist independent from speech narration, the dialogue created by the spoken elements in the game serve a key function in the overall effectiveness and pacing of *MVP Baseball 2003*. The play-by-play announcer voices are, themselves, soundmarks because they are baseball commentators in the real-world context, and are symbolic speech representations of MLB commentary. The voice of Duane Kuiper, the play-by-play announcer, is assertive and factual, while the colour commentary by Krukow is spoken in a higher-pitched, more persuasive manner. Linguistically, the words spoken by Kuiper serve to provide a detailed account of the events transpiring for the player, which could technically function to narrate the events of the game in the absence of visuals. The voices are spoken clearly, and sound as if they are sitting in the room dictating to the player, rather than being mixed into the game action. This is intended to distinguish for the player that the commentary is taking place external to the gaming environment.

In addition, the play-by-play narration can relay subtle cues to the player to take specific actions in key situations. Comments like “Change-up, low and inside for a ball” can help the player to keep track of the pitch sequence being adopted by the opposing pitcher while batting, to anticipate a pitch that might be worth swinging at. The phrase “3-2 pitch, full count. The runners have a big lead, and may be going” serves to warn the player while pitching to pay attention to runners on base. The statements by the colour commentator can be more direct and blatant in providing cues and hints to players. Krukow’s colourful comments are designed to serve a variety of functions for the players, either signaling them to adopt a particular style of take a specific action in a given scenario, or serve to reassure, congratulate, encourage, or criticize player style. For example, a blatant cue telling the player to perform a specific action will be delivered by the voice of Krukow in the form of phrases like “This is an ideal sacrifice bunt situation” or “He’s gotta focus on getting the runner home any way he can” for a player batting, and expressions like “A word of advice to the pitcher: keep the ball out of this guy’s hot zone, he could hit the ball forever”, “That curve ball just floated up there, do *not* throw another one of those” or “If they set their defense to double-play depth, they might get two [outs]” for pitching and fielding. Congratulations will be in the form of sayings such as “Great hitting” and “That was the best

---

<sup>238</sup> ©2003 Electronic Arts, Inc. *MVP Baseball 2003* (video game).

fastball we've seen him throw all game," while criticisms will be in the form of "That pitch was no where near the strike zone" or "That swing was just plain ugly" to probe or taunt the player. Reassurance is also a function provided by Krukow's speech, with phrases like "Hey, this team isn't behind by much, there's still plenty of time to get back in this game," intended to encourage the player to keep trying. The function of the Kuiper's play-by-play voice is to deliver factual information about the situational events in the game, such as calling balls and strikes or dictating the action in the field as it occurs, thereby not having too much paralinguistic variation. The play-by-play narration will take its direction from the player's actions, and will be the first sonic element following the umpire's call to accompany the current action, immediately followed by the statements of the colour commentator and the cheers or boos of the home crowd to provide the appropriate reaction of the audience. The voice of Krukow provides comic relief or colourful hints and tips, playing off the factual information delivered by Kuiper. Paralinguistic inflections are evident in the phrases of the colour commentator, projecting an obvious bias on what he considers to be a positive or negative aspect of gameplay. Interestingly, the voice of play-by-play announcer, Kuiper, functions as a linguistic element, while Krukow's colourful style has more of a paralinguistic function, creating a balance in the overall meaning of the words spoken.

The PA Announcer voice functions as a signifier of the virtual gaming environment for the player. Although the player may not be cognitively aware of the psychoacoustic impression the seemingly insignificant PA announcer voice may have on the user's psyche, it is apparent that the PA voice embodies several subtle hints about the implied spatial characteristics and depth perception, amongst other considerations, within the artificial environment. The voice legitimizes the stadium environment, in that the PA voice is filtered and compressed, to sound as if the announcements are being broadcast through the stadium's PA system. The PA announcer's voice is also processed to sound reverberant, muddled and unclear, suggesting that the voice is being heard in a large, reverberant acoustic space, at a significant, spatial distance from the player's position within the gaming environment, and contributing to a lo-fi soundscape.

Sports sims are generally devoid of gameplay music, as the aim is to mimic the audio environment from a professional sports athlete's point of view. For this reason, the majority of songs featured in the game's soundtrack are reserved for non-interactive sequences such as menu screens and replays. The songs operate to excite the player before the game or between games with adrenalizing rock music from contemporary artists. The tracks play in random order much like a jukebox, and will only repeat themselves after consecutive hours of gameplay. The music for MVP Baseball 2003 primarily functions as a motivating element that exists apart from the in-game soundscape. However, music also serves some ambient functions within the gameplay,

serving to enhance the realism of the MLB context. Organ music holds meaning for baseball fans playing the game, as does the classic stomping, clapping and chanting that accompanies the organ build-up that has been a part of professional baseball since it began over a century ago. The classic organ arpeggio that precedes the “Charge!” exclamation from the crowd serves as a soundmark in baseball stadium environment, usually serving as an instant signifier and common sound romance for baseball fans. The addition of popular music for home team player introductions is also a musical element inherent in professional sporting events. Although the game memory’s capacity cannot possibly incorporate enough songs to have a different tune accompany each home player for each team exclusive to specific stadiums, the game does provide enough variety to seem plausible. The game selects from the 25 different songs within the game, and samples one short 3-second sound bite to introduce the home player. The effect is one that gives the illusion of a vast variety of different songs being played through each stadium’s PA system, as the 3-second samples are taken from different parts of the songs –sometimes the beginning of the song, sometimes the chorus, etc.- to give the impression of a different intro for each player. The realism is maintained down the cultural details inherent in real-world Major League Baseball (MLB), in that players from foreign countries (such as Mexico and the Dominican Republic) will typically have a Latin melody as their player intro. The game mimics this by having plate appearances by players like Carlos Delgado, Roberto Alomar, Juan Gonzales, and Vladimir Guererro preceded by a sample from one of two upbeat songs with a Latin flavour, either “Mexico” by Johnny Ferreira, or “La Colegiala” by Banditos Bonitos; both contemporary Latin artists.

The energetic techno music background for the Homerun Showdown creates a frantic mood for gameplay, which suits the style of play that must be adopted in this stage. This is the only instance where music is ongoing during the course of gameplay, which changes the mood from the tense, focused approach adopted in full season play to a fun and more lenient (although fast-paced) style of play. The rapid pacing of the techno loops function to cue players to hit as many homeruns as possible to reach the objective before the opponent in order to win the challenge.

The function of SFX is to simulate the intensity of the stadium environment, as well as having crystal-clear sounds that provide the desired effect. The crack of the bat, the ball popping in the glove, the player slides, the umpire calling balls and strikes, all of these sounds serve as soundmarks; clearly identifiable and meaningful to the player outside of the gaming context. The size of the crowd is communicated through reverberation and the layering of multiple audience streams, which creates the impression of a large acoustic space, occupied by a mass of spectators.

The chanting of the home crowd is specific to the venue in which the game is being played, which provides cultural indicators about the global environment being symbolized by the simulated stadium. The chanting of specific crowds in relation to their hometown is also presumed to hold meaning to players from different cities across the US, uniting players into acoustic communities with familiar chants and songs typically sung by their hometown crowd. For example, in Montreal's Olympic Stadium, crowds will chant "Let's go Expos!", while Atlanta Braves fans in Fulton County stadium will sing the Braves' war chant in unison to encourage a home-team rally as would transpire in the real world of baseball. Fans will also chant "Over-rated" to distract and discourage star players batting for the opposing team. The voices of the vendors will also convey specific information about the geographic location of the stadium venue, and have significance pertinent within acoustic communities. For example, when the player chooses to play a game in Veteran's Stadium in Philadelphia, vendors will occasionally be heard hawking Philly Cheese-steaks, while vendors in Montreal's Olympic Stadium can be heard calling out "Poutine!" to entice hungry fans.

While there are occasional voices that can be heard above the crowd ambience of fans and vendors, they blend in with, and are considered as a part of the crowd ambience for the sake of this analysis. They are considered as SFX because they are background elements that do not stand out as signals to the player, but rather enhance the realism of the crowd ambience in the overall stadium soundscape that the player listens-in-readiness for. Technically, they have linguistic components that can be understood, but blend into the crowd as keynotes rather than sound signals for which the player would listen-in-search. The comments shouted from individual fans in the crowd or the food and drink vendors hawking their goods to hungry fans are spoken with high energy, but are low key in the overall mix, only coming through the crowd chatter occasionally. The effect is similar to that of being at a party, where many voices are speaking at the same time but our ears tend to focus attention on one voice and ignore others spoken at similar amplitudes. This phenomenon is referred to in soundscape studies as a *cocktail party effect*, which is the ability to perceive a sound from a wall of ambient background sounds heard simultaneously.<sup>239</sup> For example, amidst the cheers and boos, the voice of a vendor shouting "Burgers! I got burgers here!" will seem to project from the crowd noise, and select voices hollering at players will be heard shouting "Come on, hit the ball!" or directed at the umpire with jeers like "Are you blind? That was a strike!" will be audible through the crowd ambience.

---

<sup>239</sup> Truax, Barry. *Cocktail Party Effect*. Handbook for Acoustic Ecology CD-ROM, Cambridge Street Publishing, 1999.

### 5.4.3 MVP Baseball 2003: Interactive Audio

The intended function of the entire audio construct is to immerse the player in the gaming environment; to make the user feel a part of the action. Interactive audio elements are what draw the user into the game, and keep the player entertained for the entire nine-inning duration of the game...and hopefully for 162 games of an entire season.

Narrative speech in the form of the play-by-play and colour commentary serve as interactive elements, reacting and sounding appropriate statements to adapt to the player's style. Speech A.I. will select the appropriate speech sample and apply it to the proper gamestate for which it relates. Not only will the play-by-play comments reflect what has just transpired on the screen, but the details articulated will be specific to the exact location and velocity of the pitches thrown or balls hit, with comments like "That fastball was right down the middle of the plate" or "That slider looked too close to the velocity of his fastball." The colour commentator will then adapt with a pertinent phrase to match the sentiments of the play-by-play narration and signal a direct response to the player, in the above case, either "The pitcher needs to throw something like a sinker or a slider; something that keeps the ball down in the zone" or "Hey partner, the top of this batter's hot zone is really red, he needs to keep his pitches down low." If a spectacular catch is made in the field by the player, a slow motion replay will be triggered on the screen, and the magnitude of the catch will be amplified through dynamic and dramatic rants from the commentators, celebrating the catch as a highlight of the game. These design elements serve to reward the player for the accomplishment.

The narration will also serve to cue the player in direct relation to the state and condition of the pitcher. An element of pitching in the game is that the player must act as manager of the team, and must make educated decisions on whether a pitcher has grown tired, and needs to be replaced on the mound. The speech narration is intended to aid the player, and will adapt statements to correspond with the state of the pitcher. If the velocity of the pitches decrease, or the pitcher is struggling with his control, the narration will indicate that the pitcher needs to be switched. If a batter is not adopting the appropriate approach to a given situation, or the fielding positions are unsuitable to the action, the commentators will make a remark to signal the player to make a change. Specific comments relating to star players as continuous statements about players in key situations will also be spoken to enhance to realism of the simulated live commentary. This is a significant advance on previous sports simulation games that would insert the name of the player separately from the description of the event, which would call attention to the speech sample implementation. For example, the colour commentator will offer witty retorts about the star player in high energy situations, such as "Mike Piazza is a key factor to a Mets' victory" or

“Ray Durham plays every game like he’s going to the electric chair.”<sup>240</sup> The dialogue flows seamlessly, rarely interrupting, and mimicking the ongoing commentary that would be heard in a baseball telecast, preserving the natural flow of speech. Other voices during gameplay, such as the PA announcer voice and the voices of the umpires, will simply be triggered as physics-driven audio, corresponding to encoded parameters to play off the visual components. For example, the PA announcer will announce the name of a pinch-hitter once his statistics are shown on the screen, while umpires will call “Foul ball” if the ball falls outside of the foul lines. The designers’ intended experience is that of a fluid commentary and spoken cues that adapt to and flow along with player actions in a real-time speech simulation of the game’s events.

Music in *MVP Baseball 2003* is the least interactive component of the audio. The intended function of the music is to keep the player motivated between games and in replay sequences, and fills in gaps between at-bats, as is the case in real baseball games. The occasional organ melody will sound during a home-team rally to encourage crowd participation, which is the only real music that adapts itself to events in regular season gameplay. Even the ongoing techno music score accompanying the Homerun Showdown is non-interactive, consisting of a never-ending loop that continues until gameplay is stopped. The lack of interactivity in the music is counter-balanced by adaptive considerations in the speech and SFX of *MVP Baseball 2003*.

The interactive nature of *MVP Baseball 2003*’s SFX are perhaps its greatest accomplishment. Many of the action sounds in the game are signal-response sounds that tailor themselves to fit the visuals. Sounds will fluctuate according to input variables such as the weight, speed and state of a given player. For example, the character of Randy Johnson is a pitcher who stands 6 foot 7, weighs about 250 pounds, and is capable of throwing at about 100 mph, while Miguel Batista stands about 5’10, weighs below 190 pounds, and can barely pitch above a 90 mph fastball. The velocity of the pitches from both pitchers will sound differently when they reach the plate, and have a different frequency and timbre for the listener. When Mets’ first-baseman Mo Vaughn, who weighs about 300 pounds, slides into second base with a double, the grating sound of the slide will be lower in frequency and will sound harsh compared with Diamondbacks’ shortstop Craig Counsell, who weighs around 180 pounds. If the batter is heavier, the “whoosh” sound when he swings the bat will be deeper; if a ball is bunted, hit lightly, or popped up, the sound of the crack of the bat will be more hollow than if a batter connects squarely with a pitch and drives it to the outfield. The crack of the bats are still heard clearly in the Homerun Showdown mode of play, although the wispy sounds of the pitch whizzing by the

---

<sup>240</sup> IGN Insider: “Sound” *MVP Baseball 2003* review, IGN.com, March 11, 2003. <http://ps2.ign.com/articles/388/388870p4.html>.

batter, or the “whiff” of the bat swinging are toned down, or appropriately drowned out in the overall soundscape because of the immensity of the crowd cheering for every homerun. The atmosphere replicates that of a frenzied crowd that might be present during the All-Star game, cheering on their favourite star players as they show off their star qualities and athletic abilities. All of these variables are designed to indicate something specific to the player, to communicate meaning, and designed to evoke a different response. Sound effects in the game form an interactive relationship with the player, each feeding off one another to accomplish a task effectively. This comes back to Wiener’s theories of feedback between humans and machines in Cybernetics, as discussed in chapter one. Sound effects are designed to enhance and/or correspond with an action to provide aural feedback for the player, suitable to a given action.

The most complex sound is, without a doubt, the amalgamation of crowd ambience. The sound is actually composed of multiple layers and streams to form an ever-changing sonic mass that ebbs and flows with changes in gamestates, events and actions instigated by the player. First there is the wall of ambient chatter that flows continuously throughout the game, as would be the case in the real world; silence could never exist in the large, reverberant acoustic space of a stadium packed full of screaming fans. Different variable streams of crowd sounds overlap, each with impulsive, changeable parameters that intensify with events that augment the gamestate. The home crowd’s cheers and jeers will intensify, increasing in amplitude with every significant or subtle shift in the game. A run scoring, either for or against the home team, would inevitably incite a response from a crowd, which is replicated in the game. However, subtle shifts in atmosphere, such as the home pitcher about to deliver the two-strike pitch to the opposing batter, will also evoke a dynamic swing in crowd enthusiasm, in anticipating a possible strikeout. The crowd sounds will increase gradually with a heightened gamestate, as would happen with a real crowd, each fan feeding off the overall energy present in the cumulative mass of cheering spectators. The maximum crowd amplitude and chaotic screaming and frenzied applause would reach its peak if a home player were to hit a game-winning homerun in the bottom of the ninth inning. Replicating the real world, this would be the climax of the game, with the most adrenalizing and crowd-pleasing scenario possible, which would garner the most intensive response. The result is a mass of screaming, cheering, whistling, and clapping that borders on distortion. This maximum intensity of crowd applause is featured most frequently in the Homerun Showdown mode of play, which reaches a screaming climax that causes an excitation effect, signaling satisfaction for the player, each time a ball is tagged for a homerun. Because the screen is split in two, and the audience is divided between two players competing for the longest distance in homeruns, the crowd volume is decreased, and panning parameters will be



stereophonically specific in either speaker in relation to the player who hits the homerun; either on the right or left of the screen.

Perhaps the most dynamic and unpredictable features of the crowd that is designed to enhance interactivity and realism for the player in the virtual stadium soundscape are the group chants, heckles and shouts of approvals specific to individual players and gamestates. Individual voices in the crowd will sound sporadically for specific players batting in precise situations. For example, when a star player bats for the home team in a potential scoring situation, or heightened gamestate, individual voices emanating from the crowd such as “Come on, Robbie!” for Roberto Alomar, or “Let’s go, Barry!” for Barry Bonds, will be exclusive to the character performing the action in a given gamestate. Conversely, voices in the crowd can be heard mocking an opposing player, for example, shouting “You stink, Guerrero!” for Montreal star outfielder Vladimir Guerrero in an opposing stadium. Crowd chanting augments the realism of the home crowd supporting their team, and also serves to intensify a situation for the player batting in a key situation. For example, the crowd will chant the name of Atlanta Braves' star outfielder, Chipper Jones' first name “Chipper! Chipper! Chipper!” to support their favourite player preceding a potential hit, while the crowd will taunt an opposing player, clapping and vocalizing “Overrated! Overrated!” to distract and intimidate the adversary. These variable sound bites contribute to the engaging, interactive atmosphere maintained throughout the game in the form of the variable crowd fusion, indiscriminately inserting energetic cues and adaptive streamed responses into the game in real-time to augment the realism of the surrogate environment.

It is clear when listening and participating in the vivid gameplay experience of baseball simulation inherent in MVP Baseball 2003’s audio, that variable and adaptive sound effects in sports sim games are continually advancing, designed to immerse the player in a participatory experience. The crowd SFX and speech commentary in particular serve to provide ongoing feedback for the player throughout the gaming experience, intended to draw the player into the virtual sonic environment in order to maintain a high level of immersion until the game ends.

## 5.5 Conclusions

After conducting an in-depth study analyzing the audio components that make up the interactive aural experience for the three games studied from diverse game genres, it becomes clear that these games are at the forefront of new interactive trends in video game audio for home console gaming. Interactive audio in these games goes above and beyond the previous model of designing audio, content-to-picture, and even reverses the relationship of audio and graphics at times, having audio cues precede visuals. Although I was aware of the significant developments and advances in audio interactivity leading up to this study, I was still overwhelmed at the interactive potential and sophistication of the audio environment inherent in the games studied. Levels of interactivity for speech, music and SFX vary depending on the game genre, as stated in chapter three. The level of interactivity in each of the games analyzed conforms to the specifications particular to its game category. MVP Baseball 2003 features interactive audio ingredients in speech commentary and crowd sounds that serve to replicate the sports simulation experience, refraining from exaggerated sounds in order to enhance the realism of the gaming experience. While the baseball game lacks interactive music, fantasy sports games like SSX Tricky depend on adaptive musical elements that layer or subtract instrumental streams, designed to cue the player to take specific action. The speech and SFX are exaggerated in SSX Tricky, while the music is intended to allow for a high level of interactivity, and to provide subtle shifts to signal an evolution in play style for the user. SOCOM features suspenseful music and SFX that work together to provide subtle auditory cues to the player, and is intended to create a suspenseful atmosphere that maintains unpredictability and randomized sounds that occur in relation to the player. While the music and SFX react primarily to gamestates, the speech has a high interactive potential to engage with users, as players are required to communicate with spoken elements in the game in order to complete objectives, forming a cooperative unit with the artificial characters to advance the storyline and survive within this volatile and tension-filled surrogate environment.

The interactive speech intrinsic to all three games, that are designed to both communicate and take cues from the player, forms a dialogue between the virtual characters in the game and the player controlling them. Speech is at the forefront of games like SOCOM and MVP Baseball 2003, intended to guide the player through an interactive storyline that can adapt to the player's style and evolving gamestates. SOCOM is an example of how speech tools and systems are designed to give the player an element of control, in directing the characters verbally, and having characters vocalize pertinent feedback to acknowledge the player's commands, or to provoke desired responses to advance the plot.

Music plays less of an interactive role in MVP Baseball 2003, which is understandable given the fact that music is inherently an underlying element used sparingly in the overall audio construct of sports game simulators. Music in SOCOM serves an interactive function in enhancing the suspenseful nature of the FPS game, which undergoes dramatic shifts, designed to guide the player through gameplay. Interactive music for SSX Tricky is evidence that there are steps being taken toward a fully interactive score, in that the player remixes and builds the soundtrack and pacing through the style and progression of gameplay in this fantasy sports game.

Interactive SFX adapt themselves more to shifts in the physical environment and gamestates in FPS games like SOCOM, but are interactive in as much as they have been designed to activate in response to a player's actions initiated on the screen, and provide directional cues through binaural localization. SSX Tricky features adaptable sound effects that will conform to input variables, depending on character traits and styles of play employed by the user, proving that fantasy sports games immerse the player with engaging, over-the-top and variable SFX. MVP Baseball 2003 sets the bar for sports sim games, featuring advanced SFX tools and A.I. that serve to implement a continuous and real-time ambient soundscape that constantly ebbs and flows in relation to what the player initiates.

These games are paving the way for advances in interactive game audio, designed to allow for more realism and greater participation to enhance the immersion of the interactive gaming experience. Through audio, the player now has a greater potential to influence or even directly control the game's audio structure; customizing the soundtrack through changes in the user's style of play. While the graphics are limited to portray a select number of image variations, the audio has the capacity to dramatically and dynamically shift in response to the player, which serves to increase the overall potential for a truly immersive and thrilling participatory entertainment experience. As Aaron Marks states about designing sound for games, audio is "an integral part of any game, equal in importance to artwork, music, and gameplay. Good [audio] create[s] an impact which rounds the entire gaming experience... designed to completely absorb the player into a virtual world, making it believable, entertaining and satisfying. Continuous ambient sounds keep the player from being distracted by the "real" world, ensuring game silence doesn't ruin the immersive effect of a game."<sup>241</sup>

Having analyzed these games in depth to deconstruct the sounds specific to each game, their function within their respective surrogate audio environments, and the potential for interactivity in each audio system within the soundscape continuum, it becomes clear that audio interactivity in video games has evolved far beyond the primitive signal-response relationship of

video game audio in past years. Interactive audio considerations are garnering much deserved attention and resources to parallel sophisticated graphics, and serve to complement or even surpass the participatory capabilities of graphical components. Because listeners engage with sound and music on a subjective and emotional level, interactive audio has the power to evoke responses from players, incite reactions, serve as cuing and foreshadowing agents, and provoke participation for an ongoing immersive experience. Interactive audio intends for players to get a sense of satisfaction from the audio components, thereby having the potential to become engaged into the virtual sonic environment that encompasses them. More than just controlling a virtual character on the screen, audio is designed for users to hear and feel what it's like to be that character within the environment; listening and interacting with the encompassing soundscape's spoken, musical and ambient sonic components. The "feedback loop" that is created for the player through interactive game audio is constantly evolving to provide a unique audio experience each time the game is played, designed to work together to aid players in the environmental context of the game, either through deliberate or understated feedback.<sup>242</sup> Because of the changeability of the play experience, a high replay value of the games is probable. Because of rapidly advancing audio technology, sound designers are quickly overcoming technical limitations, and are able to expand their creative horizons to generate truly artistic, tangible and adjustable soundtracks. Excitement is building for sound artists in the video game industry, in having the ability to overcome limitations in storage capacity and financial restraints, to implement adaptive audio through sophisticated tools and algorithms, to take interactive audio to pinnacles never before imagined. In the next chapter we will discuss some designing issues inherent in interactive audio implementation, as well as discussing the future of interactive audio in home video games.

---

<sup>241</sup> Marks, Aaron. *The Complete Guide to Game Audio*. CMP Books, Kansas: 2001. p.221

<sup>242</sup> Pidkameny, Eric. "Levels of Sound." VGMusic.com, *Video Game Music Archive*: 2002. <http://www.vgmusic.com/information/vgpaper2.html>.

*Video games are pure entertainment that can stimulate our senses. The ability to provide quality sound happens to be one of the most recent advancements to game evolution. ..With these new advancements, we are discovering the need to be skilful and in control of what the player hears and feels on an emotional level. [...]Music can make you laugh, make you cry, make you aggressive, make you docile, make you fall in love, or make you hate. It is a powerful force and when used properly, can bring an entirely new dimension to a gamer's experience... By understanding the effect and exploiting the deep-rooted psychological aspects of music and sound, video games can be taken to the next level of believability and entertainment.<sup>243</sup>*

-Aaron Marks  
Video Game Audio Designer

## **6. Interactive Audio in Video Games: An Ear-Opening Experience**

As the above quote suggests, audio has the power to move people, and to engage players on multiple levels to create an immersive, participatory experience. As I have shown in this study, the systems of acoustic communication can be readily applied to video game soundscapes in a similar manner to their use in deconstructing sonic environments in the real world. The communicational systems of the continuum I have applied in this study -speech, music and soundscape- can be used to deconstruct the sonic elements within the surrogate audio environments in video games and to analyze their functional role within the virtual sonic environment. Furthermore, the sonic function of those sounds comprising the virtual sonic environment intrinsic to each game can be analyzed within the overall construct of the audio mix, to determine the role of audio interactivity in the virtual environment that encompasses the player. It becomes clear, after using a soundscape studies approach to the textual analyses of specific games adopted in this study, that audio no longer functions merely as sonic responses to support visual components, as in previous video game regimes. Audio is now a composed art that combines complexity and sophistication, as well as intricate composition and orchestration in order for the sonic components inherent in the surrogate environments to function in conjunction with players, to support an interactive encounter. Sound is now considered a primary component, moving into the foreground of game development to be paralleled with graphics, in the overall effectiveness of the interactive gaming experience, applied and orchestrated in the early design phases of game production. Chance Thomas, audio director for HUGEsound, speaks about "Day One Audio", which he denotes as the current process of the game development process, where

---

<sup>243</sup> Marks, Aaron. The Complete Guide to Game Audio. CMP Books, Kansas: 2001. p.290-291

sound professionals are brought in at the earliest conception of games to build and distil a foundation and technological framework for the interactive product. In his article “Day One Audio” for the webzine *music4games.net*, Thomas quotes art director Mark Aro of Big Sky Interactive as pinpointing game audio as the “key to the gaming experience” that “links to truly immersive game-play... For me, unique sound and music... should be implemented as early as possible.”<sup>244</sup> Thomas highlights the fundamental importance of the audio team in the design process, saying “audio should be addressed from day one” to implement interactivity, in contrast to traditional approaches of treating sound as an afterthought near the end of the development cycle.<sup>245</sup> The level of quality and sophistication of modern video game soundtracks, with complex arrangement, composition, and optimal resolution, continues to evolve, and new methods of interactive composition and digital surround localization serve to position the listener within the virtual sonic environment. The immersive nature of the participatory experience is augmented by soundtracks that adapt themselves to player actions in real-time, with audio algorithms and AI that react to a player’s style, actions and impulses.

### **6.1 Conclusions & Impressions: We’re All Ears**

Soundtracks for home video games can be analyzed as functional, virtual sonic environments, acting both as atmospheric agents and signifiers, designed to work interactively with the player for an immersive experience. Unlike earlier games, based on linear-based audio principles, a growing number of modern game audio environments can be analyzed much in the way real-world sonic environments are in the field of soundscape studies. As described in chapter two, soundscapes can be analyzed in terms of the signals, keynotes and ambient auditory characteristics that make up the sonic environments. They can also be deconstructed in as far as they provide connections and associations outside of the soundscape being studied, as is the case when examining soundmarks and acoustic communities. It is worth noting that some virtual sonic environments inherent in modern video game genres have become so sophisticated that they mirror the complex variety of real-world environments, designed to encompass the player in a virtual soundscape. In this sense, the surrogate sonic environments inherent in some games are tending toward vivid reconstruction of the variables that compose functional environments. Although it may be overly optimistic, AZ promotes the idea that sound designers might apply research around acoustic theories and soundscape parameters in order to design virtual

---

<sup>244</sup> Thomas, Chance. “Day One Audio” Published in *music4games.net*, October, 2002.  
[http://www.music4games.net/f\\_dayoneaudio.html](http://www.music4games.net/f_dayoneaudio.html)

<sup>245</sup> Ibid.

environments as real soundscapes are designed, taking into account acoustic and psychoacoustic variables, could create entirely new ways for incorporating interactive environments in video games. As theories of feedback within Cybernetics suggest, the function of game audio to provide aural responses to a player's actions creates a feedback loop that is continually evolving to provide a new and highly immersive experience for individual players each time a game is played. Author of The Algorithmic Composer, David Cope suggests that real-time digitally mediated composition creates "potential for improvisatory dialogues between human and machine," which he foresees as a positive development of real-time, algorithmic composition.<sup>246</sup> The same potential exists for the creation of virtual soundscapes, although it is unclear as to whether the game industry will be motivated to pursue this direction to the fullest extent.

The complex virtual sonic environments intricately crafted by sound artists serve to immerse the player within a fantasy context that simulates environmental cues through DSP modulation, spatial movement, textures, and even multi-channel environments with the advent of Dolby Digital Surround technology. The surrogate audio environments constructed for modern games function much the same as soundscape compositions, in that an environmental context is created by the designers, calling on the player's past experiences, associations and patterns of soundscape perception that are tactically implemented within interactive compositional strategies.<sup>247</sup> The term "soundscape composition", coined by composers in the World Soundscape Project at Simon Fraser University, denotes pieces that are composed to "enhance the listener's awareness of environmental sound", and can only be successful if the composition "changes the listener's awareness and attitudes toward the soundscape," thereby changing the listener's relationship to the real environment.<sup>248</sup> These artistic considerations are not likely to be fully explored by the game industry, however the techniques of soundscape composition and the contextual perceptions theorized in soundscape studies can be readily applied to the virtual sonic environments constructed for video games. Virtual soundscapes inherent in modern video games establish a functional context for the player, incorporating cues and environmental elements to position the listener within the virtual environment for greater immersion, and to serve as the aural foundation for the interactive experience. Combined with vivid graphics, pragmatic characters, intricate storylines and true-to-life venues, modern video games for home consoles are designed to provide a virtual, interactive entertainment experience that stimulates players on

---

<sup>246</sup> Cope, David. The Algorithmic Composer. A-R Editions, Inc. Madison, Wisconsin, 2000. p.246

<sup>247</sup> Truax, Barry. Acoustic Communication. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed., p.237

<sup>248</sup> Ibid. p.237

multiple levels, thereby immersing the user beyond that of conventional, more passive entertainment forms.

Levels of realism, as well as interactivity in audio for video games can vary according to the game and genre for which it is intended. As my study illustrates, levels of realism to replicate real-world sonic environments will be of the utmost importance in sim games like *MVP Baseball 2003*; while exaggerated sounds to create a suspenseful atmosphere will be inherent in FPS games like *SOCOM* to create dramatic tension; and fantastic environments completely eclipsing that of the real-world soundscape, as in *SSX Tricky*, will be the goal to situate the player within the exciting, exaggerated environment of fantasy sports titles. Diminished levels of realism do not necessarily encumber the immersive nature of the game, as long as the auditory structure of the game is appropriate within the context of the gaming environment, and stays consistent with the expectations of the user. Regardless of how overstated or embellished the virtual sonic environment may be, it still has the power to immerse the player, and set the stage for an interactive encounter. Levels of interactivity are necessarily varied from one game genre to another. As my study demonstrates, certain games like *SSX Tricky* will call for an interactive musical score in order to signal the player to take specific actions, or to gesture positive or negative comments on player style. Other games like *MVP Baseball 2003* will have very little interactivity inherent in the musical score, but will rely on vivid, adaptable sound effects to recreate a highly realistic and authentic gaming context. FPS games like *SOCOM* may feature music that is event-driven and SFX that corresponds to the visuals, but the act of conversing in a real-time dialogue with virtual characters in the game through interactive speech can serve to enhance the interactive nature of the participatory experience. The implications are that, although aspects of realism and interactivity fluctuate depending on the genre of the game, and elements of speech, music and sound effects can hold more emphasis in some titles than others, the widespread result is that advances in video game audio are designed to allow for more interactivity, and heightened immersion in surrogate gaming environments.

Considerations in interactive composition and implementation by sound designers serve to create a new interactive entertainment experience that captivates and immerses players into a fantasy world more than ever before. The complexities of surrogate audio environments that position the listener within the context of the virtual world can be deconstructed to reveal the sonic elements at the root of the soundtrack. Detailed analyses like those employed in this study can expose the instrumental, functional roles they play in creating a realistic, engaging experience, making players feel as if the game has been individually customized through interactive audio considerations. JMO indicates that design for game audio is shifting toward new



methods of composing and designing in order to maximize interactivity, which necessitates new ways of thinking about and approaching sound. Elements of speech, music and SFX must work together to function as more than simple signal-response triggering, but rather to create an atmosphere that can act to cue the player to take specific actions, and guide the user within the surrogate, interactive audio environment that foreshadows events and player progress as well as providing mood and context. An unpredictable, real-time soundtrack that ebbs and flows in relation to the player's actions taken in the game, fluctuating in correspondence with a player's specific style and performance and relative to changes in the artificial environment, is intended to create a unique interactive experience each time the game is played. The unpredictable nature of the gaming experience, due in large part to interactive audio considerations, has the capability to increase the replay value of the game, keeping players coming back for more in efforts to improve their gameplay. Now, with home gaming console technology empowering sound artists by giving them the necessary tools and resources to fulfill their creative visions, the experience of home gaming has reached new levels of interactivity.

## **6.2 Issues Arising from Interactive Audio Design**

While I have spent much time illustrating and documenting the advantages and merits of interactivity in video game audio, there are still issues and negative considerations to be noted. The designers admit to some complications that can arise from implementing interactive audio in video games. As JMAC explains, the video game industry is still very young and, while its evolution is exciting and allows for designers to be flexible, creative and experimental, he admits that "sometimes design efforts are doubled without achieving the desired effect in game audio." Much of interactive audio's success, he explains, depends on audio designers keeping up with advances in algorithms and tools, as sound artists are always confined to hardware limitations that constrain the designer's abilities for the console system the game is being designed for. IM explains that sound and graphics still compete for CPU time, which can be frustrating. He suggests that the ideal scenario for space preservation and sound quality would be to use MP3 files that are already compressed, although this is often not employed because the act of decoding MP3s for sample-playback in games can be quite costly. ADPCM compression strategies, while inferior to MP3s, are the cheaper alternative. IM predicts that compression and decoding CPU time will continue to be a factor designers will be forced to wrestle with, although audio technology and CPU memory capacity is quickly evolving. "Technology can be very restrictive," JMAC continues, "video games have only come into their own as a viable art form recently. It's like when movies began in the 1920s: in the early stages everyone just went at it blindly with no

standardized techniques; everyone just experimenting.” JMAC admits that design techniques are sometimes inconsistent within the audio design team, which can cause delays in the production cycle. However, he adds that methods for implementing interactive audio are becoming more formulaic, while still allowing for creativity and flexibility. The consensus from the designers interviewed in this study is that, while technological limitations can serve as an impediment to advancements in interactive audio, the limitations are becoming significantly fewer, and technological constraints are becoming less of an issue. There was an overwhelming atmosphere of optimism and excitement from sound artists at EA that interactive considerations in game audio are improving dramatically. As IM acknowledges, algorithms need to be improved to incorporate more analysis and layering in real-time audio systems. Aesthetic decisions are constantly being made on how to juxtapose sonic elements within the environments, how to make soundscapes more dynamic, and what factors must be considered to make automated responses readily available for AI systems.

Musically, JMR and JMO both recognize that interactive composition and the need for real-time transitions can be a challenge. JMR notes the problematic nature of abrupt shifts in musical transitions from one melody to the next. For JMR, shifts in music must be seamless and fluid, serving as a continuous, flowing score in synch with a player’s actions. The danger in being abrupt, he explains, is that the soundtrack becomes disjointed, calling attention to the act of playing a game, thereby taking away from the perceived realism, and detracting from the overall quality of the interactive experience. JMR describes possible methods for incorporating smooth transitions between musical segments, such as adding a drum fill, a delicate pitch shift, or ‘end events’ to serve as subtle, gestural transitions instead of abrupt, jarring movements in the score. Sometimes JMR even attempts to incorporate silence as a transition to signal something to the player rather than bombarding the user with blatant audio cues. GD also mentions subtlety as an issue designers struggle with. GD prefers that sonic cues only activate occasionally, and makes a point of integrating subtle cues and transitions, which serve as understated techniques to apply interactivity, without forcing obvious signals on the player.

Although interactivity is an overwhelming achievement in the eyes of the subjects interviewed in this study, there are still problems inherent in implementing interactivity in audio, especially with regard to licensed music. While video game audio is evolving to incorporate more interactivity, allowing for a more immersive gaming experience, the act of integrating licensed, popular music into games can be a detriment to the interactive process. JMO explains the difficulties inherent in adapting licensed music to games, in that they were not composed for the purposes of interactivity. In order to customize pop music for the purposes of structuring them

interactively in video games, designers need access to the original master tapes from the record label in order to create samples to be programmed into real-time music algorithms. There is resistance from record companies in allowing sound designers access to the original components of pop songs, because they generally do not wish for the recordings to be tampered with, and want the original format of the song to be included in the game. However, in order for music to function interactively in the context of game soundtracks, designers need to be able to reconstruct the song's components within music algorithms so that they can be applied to gameplay in real time. JMO admits that working with licensed music, and the restrictions (imposed by the record industry) that come with it, can be challenging in maintaining an interactive musical concept. He explains that it is sometimes difficult to achieve a balance of featuring licensed music in the game while adapting a song's components to an interactive construct, but without tampering with the original composition. JMO explains:

It's a bit of a trade-off, that you have to be aware of interactive concerns when licensing music... You cannot just license any music track and throw it in the game, you have to [first] judge whether a song is right for the game, then decide whether the parts of the song can be obtained, and then whether a session needs to be scheduled with the artist to re-record smaller parts of the song [for use] in the game.

While some artists and record labels can be reluctant to handing over the rights of published songs to sound designers, an agreement is usually reached whereby the song artists are compensated, and sound designers are granted the liberty to modify samples from the original source recordings to function interactively in the game, as can be heard with the RUN DMC song "It's Tricky" effectively executed in *SSX Tricky*. JMO acknowledges that the incorporation of pop music in games can have sizeable benefits to suit both the music and video game industries and can be attractive to prospective gamers, but maintains that music that is composed interactively for the purpose of being adapted in real-time during the game produces the best results in crafting a truly interactive musical score.

During the interviews with sound designers at EA, the fundamental issue that designers continue to be faced with is the ongoing problem of repetition, although responses obtained from the sound artists varied as to the extent of the problem. JMR refers to repetition as the 'nemesis' for sound designers that continues to be struggled with in interactive audio implementation. Musically, JMR explains, the score introduces a melody during which the player has a certain amount of time as a "window of opportunity to accomplish a task." If this is not done in the allotted time, he continues, the music is programmed to either fade out, or will repeat the previous melody. In contrast to Hollywood movies, JMR distinguishes "in film [the score] is

predetermined, but in games [sound designers] are at the mercy of someone pushing buttons.” This is one of the key considerations for sound artists in designing interactive audio, in that sound must function on a subconscious level with the player to cue an action or instigate a response, and must be dynamic at all costs. JMO acknowledges the challenges inherent in avoiding repetition in interactive music, but notes that repetition “has historically been something [interactive music] designers can get away with in games.” Although he acknowledges the evolution of game music from primitive looping methods that repeat sections of the score over and over again, he maintains that repetition, with respect to music, can be tolerable more than with speech elements. JMAC echoes this sentiment, conceding that repetition in speech “can kill a good game soundtrack”, especially with regard to jokes and witty comments spoken by announcers in sports titles. Speech repetition is rarely evident because console memory is programmed to remember samples that have been activated, and prioritizes un-played samples accordingly. However, “if players turn off the console between games, it takes the power out of the hands of audio designers,” and in that case, JMAC admits that repetition can result. “Kids are starting to figure it out,” JMO remarks, “and they want more variation.” JMO mentions the capabilities with current console memory advancements to incorporate more songs in games to increase variety in musical scoring, such as the game *AMPS* for the Xbox system, which he says features 300 songs. “Heck, anyone can do that,” he exclaims, “that’s not as artistic as arranging an interactive score, which is the focus [of sound designers] at EA.” JMO suggests that including hundreds of songs in a game shifts the focus to one of quantity instead of quality, and says he prefers working with 30 interactive songs that adapt to the player during gameplay, which he sees as a better alternative, using music to its full, interactive potential.

IM says that advances in audio technology for home game systems serve to make samples more manageable, thereby avoiding repetition. “It’s predetermined which samples are repeatable” IM explains, “and different levels of control are imposed to ensure that the same ones aren’t repeated back to back.” JMR claims that there is “no reason to hear loop points in this day and age with pieces [of music] that are altered in real-time, through random access memory, to change the dynamic of the melody.” Repetition can be avoided through subtle changes in the frequency content, layering and panning offsets, amongst other variations. In order to make repeated sections unnoticeable to the player, JMR says designers can “attempt to incorporate more ‘faceless’ samples into stream patterns” rather than obvious, redundant elements that the player can easily pick up on. IM maintains that continued analysis by sound artists, and refined techniques of layering will make repetition and redundancy of samples a non-issue in the near future.

### 6.3 New Relationships & Directions: Now Hear This...

While some home gamers may not have the technology to experience the full effect of the surround audio environments created for modern games, new relationships formed between Dolby Laboratories and the video game industry serve to raise the bar in game audio quality and interactivity. A recent survey commissioned by IGN.com, an online interactive entertainment news site, reports that 37 percent of all gamers polled currently play video games on a home theatre system capable of reproducing Dolby Digital 5.1 technology.<sup>249</sup> According to a SEGA spokesperson, Dolby Digital 5.1 allows for a more immersive game-play experience:

... the player's experience is made more realistic as the enemies' sounds come from nearly every direction. These audio cues increase accuracy and realism during game play... allowing [players] to anticipate key plot points or enemies approaching from different directions, which makes the experience that much more immersive.<sup>250</sup>

Dolby has also developed the Interactive Content Encoder, which enables real-time Dolby Digital encoding in games, to function in combination with Dolby Digital, Dolby Surround, and Dolby Pro Logic II, incorporated into hundreds of major game titles on all major game platforms, allowing for optimal sound quality and interactivity to maximize the auditory experience of playing video games.<sup>251</sup>

Previously, both the motion picture and music industries were hesitant to grant game designers the rights to their samples for inclusion in video games. However, the overwhelming success and superior quality of pop music and movie clips featured in games has caught the interest of the music and film industries. The fact that video-game publishers have strayed from focussing solely on violent games to target consumers may also be a factor in making the inclusion of pop music and movie clips more appealing as a marketing strategy. A variety of diverse game genres continues to emerge, offering consumers a plethora of games to choose from. Although violent FPS games continue to be some of the most popular, and are consistently the focus of much negative publicity and public scrutiny, there is a growing trend to expand on the previously limited genres of popular (and often violent) games. As chapter three of this study demonstrates, there are currently many alternatives to violent games, and even FPS games with violent undertones, such as *SOCOM: U.S. Navy SEALs*, focus on tactical strategies, covert operations, and the gathering of enemy intelligence rather than mindless killing. Video games based on popular films such as *Harry Potter* and *007 James Bond* have had such success that the

---

<sup>249</sup> Arnold, Jim. (Business Wire: Dolby Press Release) "Dolby Digital 5.1 Turns Up the Volume..." San Francisco, August 20, 2002. [http://biz.yahoo.com/bw/020820/200266\\_1.html](http://biz.yahoo.com/bw/020820/200266_1.html).

<sup>250</sup> Ibid.

<sup>251</sup> Ibid.

Hollywood movie industry has a vested interest in forming a marketing relationship with the electronic game industry to promote their movies. In fact, the recently released game *Enter the Matrix*, based on the immensely popular *Matrix* movie trilogy, introduces new ways for combining video gameplay with movie storylines, making the video game an interactive extension of the movie. The game *Enter the Matrix*, featuring characters and scenes featured in the movie *The Matrix Reloaded*, is being advertised as the “story-within-the-story” for players to take part in the plot and engage in the unfolding storyline.<sup>252</sup> The game also provides hints and foreshadowing about the upcoming movie sequel, giving gamers the edge in acquiring the newest available information about the next, highly-anticipated episode of the Trilogy. David Perry, President of Shiny Entertainment, says it’s more than a game about a movie, commenting on the unique gaming experience in the following passage.

This game isn’t just set in the *Matrix* universe – it’s an integral part of the entire *Matrix* experience, weaving in and out of the highly anticipated 2nd installment in the film trilogy, *The Matrix Reloaded*™.<sup>253</sup>

The blurred lines between the gaming experience and the movie plot are intended to make the gamer feel connected to the *Matrix* universe. The act of surviving in the game becomes something more than just a play experience, as gamers can be made to feel as if they are an intricate part of the team assembled to accomplish detailed objectives, as in the FPS game *SOCOM*.

Perhaps one of the most significant marketing developments in recent media history is the partnership of major music labels and artists with Electronic Arts, the world’s leading developer and publisher of interactive entertainment.<sup>254</sup> The recent addition of worldwide music executive, Steve Schnur, to EA’s Canadian headquarters in Burnaby, B.C. is a major step toward the new marketing relationship of the video game and music industries. In an interview with Schnur by Music4games.net, EA’s VP of music discusses the formation of “EA Trax”, a program that partners EA with every major music label, thereby providing exposure for new music artists through the active medium of video game entertainment. A prime example of the merger’s effectiveness can be seen in the success of the Capitol Records group OK Go, whose debut single “Get Over It” was introduced in *Madden NFL 2003*, a popular EA sports sim title. Music Publisher Canada reports that radio stations were initially slow to add the song to their rotation, but surges of requests from *Madden NFL 2003* fans propelled the single, which subsequently

---

<sup>252</sup> Warner Bros. Inc. © 2003. *Enter the Matrix: Game Info*, <http://www.entertthematrix.com/html/index/html>.

<sup>253</sup> *Ibid.*

<sup>254</sup> Music4games.net “EA Trax Interview: Steve Schnur” October, 2002. [http://www.music4games.net/f\\_eatrax.html](http://www.music4games.net/f_eatrax.html).

debuted at No. 1 on *Billboard's* Heatseekers chart.<sup>255</sup> Schnur explains “the [music] labels and EA will connect at retail, on tour, online, and in the game, all making the gamer’s experience a musical one. [...] If the experience becomes an extended entertainment experience, everyone wins.”<sup>256</sup> With the decrease in sales of music in recent years, the merger allows for increased exposure of new songs and artists to an ever-expanding interactive entertainment market, of which EA holds 20 percent,<sup>257</sup> providing gamers with the advantage of hearing new music first through the video-game medium. This new relationship enables consumers to rely on video games as a destination to seek out new music that will be heard in the coming months on radio and TV. EMI Music Publishing Canada President, Michael McCarty, is confident that the relationship will serve to cross-promote both the products of EA and those of major music labels, maintaining that “music doesn’t just make the gaming experience better, games have the power to help break new music.”<sup>258</sup> Schnur adds that the addition of new licensed recordings will make gamers feel like they’ve discovered the song through the game, thereby making EA games part of the label’s marketing plan. Schnur suggests that EA’s demographics are similar to those of MTV, and remains confident that gamers want familiarity and are willing to make discoveries in their gaming experience.<sup>259</sup>

As discussed in the previous section, Issues Arising from Interactive Audio Design, there is also the danger that new relationships between the music industry and the interactive entertainment industry may bring widespread advances in audio interactivity to a standstill. While rapid advances in interactive audio technology are a viable option allowing for designers to craft variable soundtracks that can be adaptable to player style and performance, there is the lingering possibility that limitations imposed by the music industry in granting the rights to incorporate their songs into games could encumber interactive implementation that would grant designers the freedom to make interactive scoring a standard in video game soundtracks. Although it remains to be seen to what extent the publishers of interactive entertainment industry will incorporate widespread interactive audio considerations, there is still the underlying concern that interactive advances may not expand to their full potential, as the inclusion of new music tracks may be increasingly attractive to the electronic entertainment industry from a corporate standpoint. The speculation that trends toward commercial control over music elements (specifically the fact that music labels are still hesitant about allowing designers to incorporate their products interactively

---

<sup>255</sup> Jones, Christopher. “Game for New Music” *Music Publisher Canada*, Fall 2002 Issue, p.4

<sup>256</sup> Music4games.net “EA Trax Interview: Steve Schnur” October, 2002. [http://www.music4games.net/f\\_eatrx.html](http://www.music4games.net/f_eatrx.html).

<sup>257</sup> Jones, Christopher. “Game for New Music” *Music Publisher Canada*, Fall 2002 Issue, p.4

<sup>258</sup> *Ibid.* p.4

<sup>259</sup> *Ibid.* p.4

to be customizable to the player) may delay the widespread evolution of interactive audio spanning the majority of video game genres.

In the Principles of Design section of this thesis, JMO was quoted as saying that the most effective game soundtracks depend on music that is composed interactively, intended for algorithmic scoring. Kurt Harland, writing for *Gamasutra*, writes about the conceptual problems of interactive scoring, in that game publishers want music in games that is pleasant to listen to, but that music for listening is inherently non-interactive. Harland, much like JMO, suggests that interactive arrangement must be approached differently from traditional composition, as a musical skeleton with many branching arms that can be heard at any moment, and must be pre-arranged to respond to gamestates within the sonic environment.<sup>260</sup> This, Harland explains, can be a difficult transition for audio designers, in that game music must be composed to perform interactively rather than as a piece of music for the sake of listening enjoyment.

The emergence of new soundtrack projects such as the video game music venture entitled “Merregon” can give rise to new waves of optimism for game music designers. Merregon assembles renowned composers from the game industry to contribute new titles on a CD compiled for the sake of listening. The unique project seeks to unite game composers and designers, in order to have their music appreciated for its own merits, external to the gaming environment. The orchestral music featured in Merregon, accompanied by a richly decorated book of images, tells a story taking place in an auditory fantasy-universe, created by the composers.<sup>261</sup> The game music society embodied in Merregon is a concept that celebrates game music, and allows the sophisticated compositions of game designers to be appreciated in musical form outside of the gaming context. Select video game soundtracks are also available for sale in music stores, such as Michael Giacchino’s *Medal of Honor* soundtrack.

With game music garnering some much deserved attention and due credit for the composed art that it has become, and with video game soundtracks now eligible for Grammy Awards,<sup>262</sup> the idea is finally catching on that video game audio has progressed far beyond the primitive beeps and electronic drones of past years. As technical limitations subside, and as sound

---

<sup>260</sup> Harland, Kurt. “Composing for Interactive Music” *Gamasutra*, February 17, 2000.

[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm)

<sup>261</sup> Boecker, Thomas. Merregon Studios © 2003. “What is Merregon?” <http://www.merregon.com/main>.

<sup>262</sup> As of the year 2000, beginning with the 42<sup>nd</sup> Grammy Award Ceremony, game scores became eligible to compete for Grammys, and be honoured as music judged as ‘stand-alone music’ on its own merits in three categories: Best Soundtrack Album for Motion Picture, Television, or Other Visual Media; Best Song for a Motion Picture, Television, or Other Visual Media; and Best Instrumental Composition for Motion Picture, Television, or Other Visual Media. The National Academy of Recording Arts & Sciences (NARAS) specifies that eligibility requirements are that the game score must be commercially available as either a separate music CD, or stored in Red Book audio format on a game CD-ROM, or as an enhanced CD (Marks, 2001). To date, there has yet to be a video game soundtrack nominated for any of the above awards.



artists continue to explore their creative potential, interactive audio in video games is certain to continue its progression to surpass linear audio concepts. Sound designers continue to lead the way in crafting virtual sonic environments that respond to players in real time, to serve as the immersive, participatory experiences they have become.

#### **6.4 Audio Interactivity in Video games: A Bright-Sounding Future**

While predictions about the future of interactive audio in video games may be speculative, recent trends and advances in the evolution of game audio, combined with the foresight of the sound designers themselves, can serve to provide valuable insight as to where interactive audio is going. As I have already alluded to in chapter three, the relatively new genre of music simulation games like SCEA's *Frequency* and Empire Interactive's *eJay* series have taken a new approach to interactive music in video games. These games, while requiring a sense of timing and rhythm from the player, necessitate no prior musical knowledge, incorporating skills of music-making through the act of remixing popular songs by pushing buttons to activate corresponding notes, thereby creating real-time music. The game lets the player construct songs using different channels to represent different components of the song, such as drums, vocals, synthesizers, bass and guitars, and subsequently allows the player to modulate the instruments using effects presets. The end result is a customized remix that can be saved by players, and shared online, thus creating a global electroacoustic community. Music creation game software *House eJay*, Empire Interactive's most recent achievement, provides a built-in interface for music creation. The game requires more of a musical engineering inclination from the player, in that it doubles as a basic multi-tracking software, but functions to introduce basic sound manipulation concepts to aspiring musicians. The game contains a library of 3000 sound samples, a Groove Generator that sequences drum patterns, a stereo enhancer, a phase vocoder, various sound manipulations built into its FX studio (including an exclusive time-stretcher), and boasts CD-quality, 16-bit, 44.1kHz digital sound. The game also makes 24 tracks available to the player, and allows gamers to save their mixes in WAV or MP3 format.<sup>263</sup> Such music creation games may be too advanced for the novice music sim gamer, but they are still an example that interactivity in music games is evolving to allow players the ability to adapt the soundtrack to their specifications, giving the player more control over the auditory experience. EA's newest games, such as the fantasy sports title *NBA Street Vol. 2*, offer players access (within the options menu) to switch between "interactive" or "sequential" music playback, which gives the player the choice

---

<sup>263</sup> Hyde-Smith, Alex. "House eJay" *Music4games.net*. October 16, 2002. [http://www.music4games.net/f\\_ejay\\_house.html](http://www.music4games.net/f_ejay_house.html).

of specific songs from the 'EA Trax' songlist that the game will play, and in what order.<sup>264</sup> The term "interactive" may be an exaggerated term for this simplistic level of choice, but it is still a step toward giving the player more control over the musical experience provided in the game.

Arising from these developments is the discussion of the varying levels of interactivity offered by modern games. As video-game sound becomes more sophisticated, providing adaptable, real-time, CD-quality surround audio environments, interactive considerations in audio are revolutionizing the immersive nature of the gaming experience. It can be said that it is no longer a question of whether or not games are progressing toward full interactivity, but is rather a question of how soon it will come to fruition.<sup>265</sup> The answer largely depends on the desires of consumers. As will be discussed in the subsequent chapter, it is debatable whether gamers want full control over their auditory gaming experience. Adding to this discussion, one of the biggest names in video-game audio design, Darryl Duncan of GameBeat Inc., spoke of the future of audio interactivity in video games during an interview with Andrew Clark for Gamasutra in May of 2001. Duncan predicts that audio software packages "will allow game developers to put custom, original music in their game – not needing musicians or sound designers at all."<sup>266</sup> This statement, implying that advances in interactive audio tools will eventually make sound artists obsolete, was met with heated objection by the designers interviewed in this study. GD maintains that players (as well as developers) want a preset audio design; an operational, auditory world that has been created for them that allows for navigation and a gradual formation of the interactive experience within the simulated environment. He insists "gamers wouldn't want to create an audio design from scratch, they want built-in audio cues to give them control over their own destiny [in the game]." JMAC adds that sound design for interactive entertainment should be considered an art form that will always necessitate sound artists to design, because players still want some form of passive entertainment in addition to having an element of control. "People don't want to design their own sound," he exclaims "you have to give them the tools and creative construct before they will want to engage in the experience. There will always be a need for sound designers, just like there will always be a need for artists." JMAC predicts that a wide range of audio interactivity in games will emerge as interactive entertainment expands their scope to include a broader target audience. He speculates that while some games may focus on "a full-on, immersive cyber-experience", games providing a passive experience, but with an element of interactive control, will probably be the most popular.

---

<sup>264</sup> EA BIGSports ©2003. *NBA Street Vol. 2* (video game)

<sup>265</sup> Pidkameny, Eric. "Levels of Sound" *VGMusic.com*, 2002. <http://www.vgmusic.com/information/vgpaper2.html>.

<sup>266</sup> Clark, Andrew. "An Interview with Darryl Duncan". *Gamasutra*, May, 2001. [http://www.gamasutra.com/resource\\_guide/20010515/marks\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/marks_01.htm).

When asked about the future of audio interactivity in video games, the sound artists had numerous enthusiastic responses. JMR feels that they have already reached the maximum of what can be achieved with sample-based triggering methods. “Making software that generates audio on the fly that is malleable,” should continue to be the focus of interactive audio according to JMR, and furthermore “to improve computer intelligence to make educated decisions and real-time analysis of game sound is the next step.” JMO looks forward to developments that will make all games featuring 5.1 Dolby Surround sound a reality, while also allowing for interactive, real-time scoring that adapts to gameplay as instigated by the player. He also suggests that interactive, real-time filters could be capable of applying far more complex DSP effects, and perhaps even real-time synthesis, generating music to coincide directly with actions on the screen. JMO hopes that artists will begin to compose songs to suit multiple channels and support interactivity. Regardless, he is confident that techniques of music layering, as in *SSX Tricky*, will be used more in future game audio designs to push the boundaries of interactivity, and allow for a truly engaging, participatory musical experience. IM suggests that more analysis and layering techniques need to be researched in order to program interactive speech into a thought process that is encoded as a “stream of utterance that can be flexible and respond to inputs... the way human speech functions, as the thought process evolves.” IM maintains that speech within video game audio continues to be a key area of focus, and that improving interactive considerations for speech will make characters truly adaptive and responsive to their environment, to game events, and to other characters in the game. JMAC envisions that a program capable of simulating real-time speech synthesis patterns would function to heighten the realism of the gaming experience, as would full audio quality, eliminating the need for compression or downsizing of samples. He says that game audio has the capability to surpass the linear audio inherent in movies, and sees this as an inevitable succession because people will come to demand an interactive entertainment experience. Developments in audio will advance in leaps and bounds, JMAC foresees, once more resources are devoted to game audio, which will happen the more video games are accepted publicly as viable entertainment art forms instead of merely as ‘games’. GD anticipates more subtlety in interactive audio cues and transitions, guiding the player through understated sonic shifts. Ideally, AZ looks to the future with hopes that in-home theatres will become cheaper and more widespread, so that designers could incorporate more sophisticated audio without having to appeal to the lowest technological denominator. He also calls on more training and technical expertise for aspiring sound artists to be educated in “the interactive aesthetics of sound.” Only highly skilled designers, he attests, are capable of collaborating to create interactive audio environments that function as living entities.

In conclusion, the interactive nature of video game audio, which takes home entertainment to an entirely new level, allows audio components to adapt and conform to random gameplay actions taken by the individual player. This enhanced auditory practicality adds to the realism of modern games, and is designed to augment video-game playing into a fun, interactive, immersing and rewarding experience. One can only wonder about what the future holds for audio in interactive entertainment media, as audio technology continues to evolve quickly and steadily, reinventing itself with each generation of home console system that succeeds its predecessor. Although the designers at EA, respecting the non-disclosure agreement, declined to comment on what we can expect to see leading up to the release of Sony's *PlayStation 3* over the next couple of years, they hinted that it will exceed our wildest expectations, and take audio to new heights. As I have discovered time and time again during the research and documentation of this study, attempting to adequately detail the advances in the emerging area of interactive audio for the thriving medium of video game entertainment becomes a never-ending task, as the newest trends and developments become obsolete almost as quickly as one can document them on paper. As Steven Kent, author of The Ultimate History of Video Games concedes, writing about the interactive entertainment industry is a "game that never ends."<sup>267</sup> I have attempted to provide some insight and raise awareness about the implementation of interactive audio for home gaming consoles, and sincerely hope that my thesis serves to shed light on the emerging area of study into audio, an area that has historically been underappreciated and taken for granted, often disregarded as 'background noise'. This study could serve to validate further research into the area of interactive audio in video games. Concrete results could be obtained as to the effects of interactive audio perception, although this would necessitate longitudinal studies of the relationships formed between the users engaging in the virtual sonic environments inherent in interactive gaming. Long-term studies of this nature may allow researchers to make more concrete assumptions about the relationship between the construct of interactive video-game audio and perceptions by the players, themselves. By combining the approaches adopted in this study, coupled with theories of audience perception and psychology, one could formulate hypotheses about the sender-receiver relationship intended by interactive audio considerations within the video game medium. I would encourage my colleagues to attempt such a rewarding undertaking, although it would require much passion and perseverance. The future of interactive audio for home video game console systems is sure to undergo many exciting technological developments in the upcoming years, to allow for greater immersion within the interactive entertainment experience.

---

<sup>267</sup> Kent, Steven L. The Ultimate History of Video Games. Prima, Random House: New York, 2001. p.591

## **APPENDIX A**

# ***SOCOM U.S. Navy Seals Sound Analysis Chart***

The following chart is a detailed analysis of sounds, and their implied meanings perceived by players, as documented chronologically as they occur in evolving gamestates during gameplay in *SOCOM U.S. Navy SEALS*. The chart is structured to give the reader a detailed account of how sounds function within the construct of the continuous gameplay inherent in the *SOCOM FPS* game, and assumes the sound mix has been unchanged from the audio default settings. The path and order of the events within the documented Stages (1, 2 & 3) can occur in different orders depending on the players' adopted styles. Sounds will also vary according to the level of difficulty selected by the player prior to gameplay (varying the level of guidance provided by speech components, primarily) to suit the gamer's personal preference. The chart is meant to serve as an accompaniment to section 4.2 of the thesis, and provides the reader with a detailed account of the sounds encountered in chronological order of when they occur in standard gameplay of *SOCOM U.S. Navy SEALS*.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Opening Sequence N.I.S.	N/A	Main Theme (triumphant)	<ul style="list-style-type: none"> <li>◆ Helicopter</li> <li>◆ Gunshots</li> <li>◆ Explosions</li> </ul>	<p>Excitation effect. Excites player for gameplay to follow</p> <ul style="list-style-type: none"> <li>◆ Lead into action</li> <li>◆ Establish context</li> <li>◆ Set the stage for action</li> </ul>	Clips from different levels, like movie preview
N.I.S. Intro to Stage 1	N/A	Main Theme (abbr. version)	<ul style="list-style-type: none"> <li>◆ Wind</li> <li>◆ Explosion</li> <li>◆ Snowmobile</li> </ul>	<ul style="list-style-type: none"> <li>◆ SEALs on snowmobiles, driving away from an explosion</li> </ul>	
Mission Briefing Menu Screen	<ul style="list-style-type: none"> <li>◆ "HQ" Voice (female) gives information, mission details</li> </ul>	Main Theme (same as Opening Theme, one section looped)	N/A	Exciting, waiting for the navy characters to be "deployed"	Menu Screen
Stage 1: "Death at Sea" N.I.S. EXT. Underwater, Boat	<ul style="list-style-type: none"> <li>◆ "HQ" Voice (female) gives details of mission.</li> <li>◆ "HQ" Voice (male) reviews specific objectives &amp; strategy</li> </ul>	Suspense Theme (minor key variation of main theme)	<ul style="list-style-type: none"> <li>◆ Bubbles underwater</li> <li>◆ Waves</li> <li>◆ Rain</li> <li>◆ Boat Fans</li> </ul>	Suspenseful, dark & eerie. Danger lurking.	SEALs infiltrate terrorist ship underwater, prepare to gather intelligence on terrorist activity undetected.
Stage 1: "Death at Sea" Gameplay EXT. Boat, Rainy	<ul style="list-style-type: none"> <li>◆ "HQ" Voice (female) in headset warns of danger; offers tips for successful completion of objectives.</li> <li>◆ Russian Enemies talking, can be overheard when approached.</li> <li>◆ Boomer (Able)</li> </ul>	Suspense Theme (fades in and out with periods of silence) <ul style="list-style-type: none"> <li>◆ Drumrolls</li> <li>◆ Suspense nuances (drones with strings)</li> </ul>	<ul style="list-style-type: none"> <li>◆ Footsteps (on metal surface)</li> <li>◆ Metal clanking of ships parts in the breeze</li> <li>◆ Engine sounds</li> </ul>	Suspenseful, eerie. Danger lurking. Tension-building. Proceed with caution, stay hidden.	P.O.V. of Kahuna (FPS) on boat, crouching down and ready for action. Kahuna character moves about ship: exploring, shooting, planting devices & gathering enemy intelligence.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Stage 1: "Death at Sea" Gameplay EXT. Boat, Rainy (cont'd)	<ul style="list-style-type: none"> <li>◆ Element</li> <li>◆ Jester (Bravo Element)</li> <li>◆ Spectre (Bravo Element)</li> </ul>				
Stage 1: "Death at Sea" Gameplay EXT. Boat, Rainy	<ul style="list-style-type: none"> <li>◆ HQ Voice (female) in ear-piece interjects repeatedly to guide "novice" player through objectives.</li> <li>◆ 1<sup>st</sup> player (Kahuna) issues commands in headset mic.</li> <li>◆ Bravo Unit troops acknowledge 1<sup>st</sup> player commands.</li> </ul>	<p>Suspense Theme (fades in and out with periods of silence)</p> <ul style="list-style-type: none"> <li>◆ Reward Music activates when objective is reached (triggered by game event).</li> <li>Suspense theme resumes after Reward Jingle.</li> </ul>	<ul style="list-style-type: none"> <li>◆ Gunshots (enemy rifles distinguished by sharp attack)</li> <li>◆ Gunshots (friendly fire distinguished by softer-sounding silenced rifle bursts)</li> </ul>	<p>Suspenseful, eerie. Danger lurking. Tension-building. Proceed with caution, stay hidden. Sense of satisfaction at completing objectives, sense of urgency to complete mission.</p>	<p>Kahuna (FPS) character moves about ship: exploring, shooting, planting devices &amp; gathering enemy intelligence.</p>
Stage 1: "Death at Sea" Gameplay EXT. Boat, Rainy Game Event: Enemy Attack sequence	<ul style="list-style-type: none"> <li>◆ Russian enemies speak with urgency, expressing alarm at SEAL team's presence; at being attacked.</li> <li>◆ SEAL team (Bravo unit) will alert player 1 (in headset) of "Tango" (enemy) presence; will inform player 1 of any</li> </ul>	<ul style="list-style-type: none"> <li>◆ High-intensity, fast-tempo "panic" music plays as high-impact game event transpires; shots fired at player 1.</li> <li>◆ Short bursts of energetic theme alarms player, and instigates action.</li> </ul>	<ul style="list-style-type: none"> <li>◆ Enemy fire: loud, sharp attack of machine guns.</li> <li>◆ Bullets ricochet of metal surfaces.</li> </ul>	<p>Startling! The audio varies in intensity according to the proximity of shots fired. Directional SFX correspond to stereophonic localization to give player 1 clues as to self-location and enemies' position</p>	<p>Kahuna (FPS) character shooting, running through different parts of the ship's deck to avoid or seek out enemies.</p>



# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
(cont'd)	enemy kills or SEAL casualties.			within the sonic environment.	
Stage 1: "Death at Sea" Gameplay INT. Boat Control Room	<ul style="list-style-type: none"> <li>◆ Enemies cry out when they spot player 1. Voices sound louder in INT ambient context (reverberation &amp; absorption).</li> </ul>	<ul style="list-style-type: none"> <li>◆ Music gradually intensifies, interactively, when Player 1 fires shots.</li> </ul>	<ul style="list-style-type: none"> <li>◆ Muffled rain sounds can be heard outside.</li> <li>◆ Beeps of radar beacon</li> <li>◆ Windshield wipers outside control room window.</li> </ul>	Characteristics of an indoor space are conveyed through the acoustic design.	Kahuna (FPS) character moves about control room (interior of ship): exploring, shooting, planting devices & gathering enemy intelligence.
Stage 1: "Death at Sea", Game Event: Completion of all Mission Objectives	<ul style="list-style-type: none"> <li>◆ HQ female voice (in headset) congratulates Player 1 on a job well done.</li> </ul>	<ul style="list-style-type: none"> <li>◆ Triumphant reward fanfare.</li> </ul>	N/A	Player 1 feels a sense of accomplishments. There is a moment of celebration and a sense of achievement before the next mission is introduced.	Screen zooms out from Kahuna character's final position upon completion of the level.
Mission Details Screen (when objectives are completed successfully)	N/A	<ul style="list-style-type: none"> <li>◆ Militant, jubilant movement (major scale). 30-second fanfare loops until player pushes a button to advance to next screen.</li> </ul>	N/A	Celebration, Reward.	"Mission Details" screen gives score and overall letter grade for player 1 style and objectives completed.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Mission Details Screen (when player 1 dies, fails to complete mission)	N/A	<ul style="list-style-type: none"> <li>Severe, negative "dread" music (minor key). 30-second "tragic" music loops.</li> </ul>	N/A	Scorn & disdain reflected through music, but still upbeat and rhythmic: urges the user to try again and be persistent.	"Mission Details" screen reads "Unsatisfactory"; overall letter grade for player 1 style and objectives to motivate player.
Mission Briefing Menu Screen	<ul style="list-style-type: none"> <li>HQ (female voice) explains details of mission when prompted.</li> </ul>	<ul style="list-style-type: none"> <li>Main Theme (same as Opening Theme, one section looped)</li> </ul>	N/A	Exciting, waiting for the navy characters to be "deployed"	Menu Screen
N.I.S. (mini-movie) sets the stage for Stage 2: "Ghost Town"	N/A	<ul style="list-style-type: none"> <li>Triumphant orchestral melody (different from Stage 1)</li> </ul>	<ul style="list-style-type: none"> <li>Heavy, swirling winds</li> <li>Plane engine</li> </ul>	Building excitement. Sets the stage for the gameplay to follow.	Fighter plane flying over enemy territory. SEALs jump out with parachutes.
Stage 2: "Ghost Town" Gameplay: EXT. Snowy Alaska	<ul style="list-style-type: none"> <li>HQ Female voice (headset) provides info on new objectives</li> </ul>	<ul style="list-style-type: none"> <li>Suspense music fades in and out.</li> <li>Drumrolls</li> <li>Drones (with strings)</li> <li>Silences</li> <li>String section (activates with gamestate shift)</li> </ul>	<ul style="list-style-type: none"> <li>Light winds</li> <li>Footsteps in snow</li> <li>Crow screeches periodically</li> <li>Wolf howls periodically</li> </ul>	Suspenseful & eerie. Foreign, unfamiliar territory: Danger lurking.	Kahuna (FPS) character moves about snowy landscape: exploring, shooting, planting devices & gathering enemy intelligence.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
(cont'd)		<ul style="list-style-type: none"> <li>Staccato Cello plucking with Piano arpeggio</li> </ul>			
When using Scope of Sniper rifle to zoom in on unsuspecting "Tango" enemy	N/A	N/A	<ul style="list-style-type: none"> <li>Heartbeat of Kahuna (1<sup>st</sup> person P.O.V.), periodic, rhythmic.</li> </ul>	Symbolizes concentration, focus & the need for precision	Circular view of enemy through zoomed-in sniper lens
Stage 2: "Ghost Town" Gameplay: INT. Cabin #1	N/A	<ul style="list-style-type: none"> <li>Personal radio in cabin plays a Russian pop song (loops entire song). Radio will shut off if shot by player 1.</li> </ul>	<ul style="list-style-type: none"> <li>Radio static</li> <li>Gunshots amplified in interior space.</li> <li>Fire crackle</li> </ul>	Realism of context and culture represented through foreign language and culturally-specific music.	Kahuna (FPS) character moves about log cabin: exploring, shooting & gathering enemy intelligence.
Stage 2: "Ghost Town" Game Event: When enemy sneaks up, attacking Kahuna from behind	<ul style="list-style-type: none"> <li>SEAL team informs player 1 if Tangos are spotted.</li> <li>Enemies utter a war-cry when attacking player.</li> </ul>	<ul style="list-style-type: none"> <li>Intense, rhythmic, pulsating song</li> </ul>	<ul style="list-style-type: none"> <li>Enemy footsteps</li> </ul>	Alarming music (less intense than "panic" music of Enemy Attack sequence) intensifies dynamic mood, signals player to action, gives player a moment to react	N/A (enemies are <i>heard</i> rather than seen, approaching from behind)

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Stage 2: "Ghost Town" Game Event: When Kahuna sneaks up on an enemy from behind	<ul style="list-style-type: none"> <li>Russian enemy voices can be overheard talking when player 1 sneaks up on them.</li> <li>Enemy says "Huh?" if Kahuna is seen</li> </ul>	N/A	<ul style="list-style-type: none"> <li>Kahuna's footsteps</li> </ul>	Quiet, strategic attack from behind	Enemy characters gradually get closer as Kahuna approaches
Stage 2: "Ghost Town" Gameplay: INT. Cabin #2	N/A	<ul style="list-style-type: none"> <li>Personal radio in cabin plays a Russian pop song (different song from Cabin #1). Radio will shut off if shot by player 1.</li> </ul>	<ul style="list-style-type: none"> <li>Radio static</li> <li>Gunshots amplified in interior space.</li> <li>Fire crackle</li> </ul>	Realism of context and culture represented through foreign language and culturally-specific music.	Kahuna (FPS) character moves about log cabin: exploring, shooting & gathering enemy intelligence.
Stage 2: "Ghost Town" Game Event: Exploding Weapons Cache	N/A	N/A	<ul style="list-style-type: none"> <li>Large explosions as C4 detonations occur</li> </ul>	Satisfaction at having destroyed terrorist weapons cache	Impressive explosions as crates holding weapons are destroyed
Stage 2: "Ghost Town" Game Event: Completion of all Mission Objectives	<ul style="list-style-type: none"> <li>HQ female voice (in headset) congratulates Player 1 on a job well done.</li> </ul>	<ul style="list-style-type: none"> <li>Triumphant reward fanfare.</li> </ul>	N/A	Player 1 feels a sense of accomplishments. There is a moment of celebration and a sense of achievement before the next mission is introduced.	Screen zooms out from Kahuna character's final position upon completion of the level.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Mission Details Screen (when objectives are completed successfully)	N/A	<ul style="list-style-type: none"> <li>Militant, jubilant movement (major scale). 30-second fanfare loops until player pushes a button to advance to next screen.</li> </ul>	N/A	Celebration, Reward.	"Mission Details" screen gives score and overall letter grade for player 1 style and objectives completed.
Mission Details Screen (when player 1 dies, fails to complete mission)	N/A	<ul style="list-style-type: none"> <li>Severe, negative "dread" music (minor key). 30-second "tragic" music loops.</li> </ul>	N/A	Scorn & disdain reflected through music, but still upbeat and rhythmic: urges the user to try again and be persistent.	"Mission Details" screen reads "Unsatisfactory"; overall letter grade for player 1 style and objectives to motivate player.
Mission Briefing Menu Screen	<ul style="list-style-type: none"> <li>HQ (female voice) explains details of mission when prompted.</li> </ul>	<ul style="list-style-type: none"> <li>Main Theme (same as Opening Theme, one section looped)</li> </ul>	N/A	Exciting, waiting for the navy characters to be "deployed"	Menu Screen
N.I.S. (mini-movie) sets the stage for Stage 3: "Oil Platform Takedown"	N/A	<ul style="list-style-type: none"> <li>Triumphant orchestral melody (different from Stages 1&amp;2)</li> </ul>	<ul style="list-style-type: none"> <li>Helicopter</li> </ul>	Building excitement. Sets the stage for the gameplay to follow.	Helicopter flying over enemy territory to drop off SEAL team.

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape	Graphic/Visual
<p>Stage 3: "Oil Platform Takedown" Gameplay: EXT. Cargo ship, night (Time limit)</p>	<ul style="list-style-type: none"> <li>◆ HQ Female voice explains mission objectives.</li> <li>◆ HQ Male voice gives tactical strategy on bomb diffusion and hostage rescue.</li> <li>◆ Russian sentry guards overheard reporting periodically that the area is clear.</li> </ul>	<p>Suspense Theme (fades in and out with periods of silence)</p> <ul style="list-style-type: none"> <li>◆ Drumrolls</li> <li>◆ Suspense nuances (drones with strings)</li> </ul>	<ul style="list-style-type: none"> <li>◆ Waves</li> <li>◆ Wind (brisk &amp; swirling)</li> <li>◆ Footsteps (on cargo ship platform)</li> </ul>	<p>Eerie, covert, suspenseful. Remain undetected at all costs. Avoid being seen by sentry guards, and move quickly!</p>	<p>Kahuna (FPS) character moves about cargo ship: exploring, seeking out bombs &amp; hostages.</p>
<p>Stage 3: "Oil Platform Takedown" Game Event: When locating/diffusing bomb</p>	<ul style="list-style-type: none"> <li>◆ When bomb is diffused, HQ Male voice reports that the bombs have been successfully neutralized.</li> </ul>	<ul style="list-style-type: none"> <li>◆ Reward music (different from stages 1 &amp; 2, like relief more than militant fanfare)</li> </ul>	<ul style="list-style-type: none"> <li>◆ Beeping as clock ticks</li> <li>◆ Buzz of oil generator</li> </ul>	<ul style="list-style-type: none"> <li>◆ Signals sense of urgency for the player to act quickly.</li> <li>◆ Announces victory when bombs are successfully neutralized.</li> </ul>	<p>Fuse box &amp; digital clock facing of bombs</p>
<p>Stage 3: "Oil Platform Takedown" Gameplay INT. Cargo Ship</p>	<p>N/A</p>	<p>N/A</p>	<p>Silence (no wind)</p>	<p>Indoor context, sheltered from windy weather outside</p>	<p>Kahuna explores interior of Cargo ship for hostages</p>

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Stage 3: "Oil Platform Takedown" Gamestate: When exiting INT corridor to go outside	N/A	N/A	<ul style="list-style-type: none"> <li>◆ Suction wind gust replicates the ears adjusting to the external environment</li> </ul>	Simulates how ears adjust to different indoor-outdoor contexts. Gives impression of bad weather, makes player aware of changes in context & climate.	Kahuna exits through cabin door to exterior of cargo ship's deck
Game event (Stages 1, 2 & 3): when a grenade goes off in relatively close proximity to 1 <sup>st</sup> player character, Kahuna	N/A	N/A	<ul style="list-style-type: none"> <li>◆ Short blast</li> <li>◆ Instant dampening of all sounds</li> <li>◆ High-pitched squeal</li> </ul>	Replicates the effect of "TTS" in simulating the ringing in one's ears when a loud, sudden sound occurs.	Grenade explodes
Stage 3: "Oil Platform Takedown" Game Event: Bomb Location	N/A	N/A	<ul style="list-style-type: none"> <li>◆ Beeping sound gets louder/softer depending on Kahuna's proximity to the bomb (also pans L&amp;R)</li> </ul>	Beeping SFX serve to cue the player as to the sonic object's location, even when it is not yet visible on screen	Green flashing light (synchronized periodically with beeps) gives visual indication if once the bomb is located

# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Stage 3: "Oil Platform Takedown" Game Event: Completion of all Mission Objectives	<ul style="list-style-type: none"> <li>HQ female voice (in headset) congratulates Player 1 on a job well done.</li> </ul>	<ul style="list-style-type: none"> <li>Triumphant reward fanfare.</li> </ul>	<ul style="list-style-type: none"> <li>Helicopter returns to collect SEAL team to bring them back to Navy Headquarters</li> </ul>	<p>Player 1 feels a sense of accomplishment. There is a moment of celebration and a sense of achievement before the next mission is introduced.</p>	<p>Screen zooms out from Helicopter's final position upon completion of the level.</p>
Stage 3: "Oil Platform Takedown" Game Event: Failure to diffuse bombs within the allotted time	<ul style="list-style-type: none"> <li>HQ Female voice shouts (headset) "Abort Mission! The Terrorist leader has detonated the bombs!"</li> </ul>	N/A	<ul style="list-style-type: none"> <li>Explosions (numerous)</li> </ul>	<p>Failure to diffuse bombs and rescue hostages</p>	<p>The cargo ship erupts with a spectacular series of explosions</p>
Stage 3: "Oil Platform Takedown" Game Event: Hostages are rescued	<ul style="list-style-type: none"> <li>Husky, male voices of American hostages thank Kahuna for saving them.</li> <li>HQ Female voice congratulates player 1 on completing an objective, reminds that the objection to diffuse bombs is still pending.</li> </ul>	<ul style="list-style-type: none"> <li>Reward Music (brief jingle)</li> </ul>	N/A	<p>Reward.</p>	<p>N.I.S. of hostages celebrating their freedom is shown on screen.</p>



# SOCOM U.S. Navy Seals Sound Analysis Chart

Stage & Context	Speech	Music	SFX	Soundscape Percept/meaning	Graphic/Visual
Game Event (Stages 1, 2 & 3): Kahuna is killed	<ul style="list-style-type: none"> <li>Boomer's voice can be heard in headset: "HQ, commander is down, please advise"</li> </ul>	N/A (silence)	<ul style="list-style-type: none"> <li>"Thud" as Kahuna collapses.</li> <li>Wind swirling</li> </ul>	Increased silence indicates the pause after player 1 has been killed	Kahuna lies motionless, virtual camera zooms out and rotates around Kahuna
Mission Details Screen (when player 1 dies, fails to complete mission)	N/A	<ul style="list-style-type: none"> <li>Severe, negative "dread" music (minor key). 30-second "tragic" music loops.</li> </ul>	N/A	Scorn & disdain reflected through music, but still upbeat and rhythmic: urges the user to try again and be persistent.	"Mission Details" screen reads "Unsatisfactory"; overall letter grade for player 1 style and objectives to motivate player.
Mission Details Screen (when objectives are completed successfully)	N/A	<ul style="list-style-type: none"> <li>Militant, jubilant movement (major scale). 30-second fanfare loops until player pushes a button to advance to next screen.</li> </ul>	N/A	Celebration, Reward.	"Mission Details" screen gives score and overall letter grade for player 1 style and objectives completed.

## **APPENDIX B**

# ***SOCOM U.S. Navy Seals : Sound Count Sheet***

## **Game Description:**

*The United States Navy SEAL Team is legendary for its prowess in the art of modern warfare [...] Navy SEALs (Sea, Air, Land) take their name from the environments in and from which they operate [...] Teamwork, perseverance and exceptional weapons knowledge make SEAL teams the most lethal combat forces operating today. SEAL training also pays off in low combat casualties...As an elite SEAL commander, your team has been called in to fight terrorists in 12 of the most deadly missions ever. With more than 30 weapons, superior intelligence, intense training, and unparalleled skills and tactical maneuvers at your disposal, protect freedom wherever it is in danger.*

(©2002 Sony Computer Entertainment America Inc. SOCOM U.S. Navy SEALs Instruction Manual, p.3, 47)

Below is the list of sounds documented during the analysis of *SOCOM U.S. Navy SEALs*. The sounds were isolated and catalogued for the purpose of deconstructing the sound components in the textual analysis. This list does not attempt to accurately state all of the sound samples employed by the designers, but rather seeks to illuminate the key sounds that are perceived by the player and pertinent within the overall construct of the virtual sonic gaming environment for the purposes of this analysis. Each “stage” represents a sequential level in the game as the player progresses through gameplay.

## **Stage 1**

### **Stage 1: “Death at Sea”**

#### **SFX:**

- 1) Helicopter
- 2) Gunshots (various)
- 3) Explosions (various)
- 4) Waves
- 5) Wind (light breeze)
- 6) Bubbles Underwater
- 7) Rain (variable)
- 8) Boat fans
- 9) Footsteps (various)
- 10) Metal clanking (boat)
- 11) Engine sounds
- 12) Bullets Ricocheting off surfaces (various)
- 13) Beeps of radar beacon
- 14) Windshield wipers
- 15) Heartbeat (P.O.V.)
- 16) “Thump” ground impact when a soldier collapses

(continued)

# **SOCOM U.S. Navy Seals : Sound Count Sheet**

## **Stage 1: "Death at Sea"**

### **Speech:**

- 17) "HQ" voice #1 (female)
- 18) "HQ" voice #2 (male)
- 19) Kahuna (Able unit, first player character)
- 20) Boomer (Able unit)
- 21) Jester (Bravo unit)
- 22) Spectre (Bravo unit)
- 23) Enemy voice #1
- 24) Enemy voice #2
- 25) Enemy voice #3
- 26) Enemy voice #4
- 27) Enemy voice #5

### **Music:**

- 28) Main Theme, Major & Minor scale (multiple variations & durations)
- 29) Suspense Drumfill
- 30) Suspense Drones
- 31) Suspense Strings
- 32) Reward Fanfare

## **Stage 2**

### **Stage 2: "Ghost Town"**

- 1) Wind (heavy, swirling; variable)
- 2) Plane engine
- 3) Helicopter
- 4) Footsteps (various)
- 5) Gunshots (various)
- 6) Explosions (various)
- 7) Crow screeching
- 8) Wolf howling
- 9) Heartbeat (P.O.V.)
- 10) Radio static
- 11) Fire crackle
- 12) Snowmobile
- 13) Bullets Ricocheting off surfaces (various)
- 14) "Thump" ground impact when a soldier collapses

(continued)

# **SOCOM U.S. Navy Seals : Sound Count Sheet**

## **Stage 2: "Ghost Town"**

### **Speech:**

- 15) "HQ" voice #1 (female)
- 16) "HQ" voice #2 (male)
- 17) Kahuna (Able unit, first player character)
- 18) Boomer (Able unit)
- 19) Jester (Bravo unit)
- 20) Spectre (Bravo unit)
- 21) Enemy voice #1
- 22) Enemy voice #2
- 23) Enemy voice #3
- 24) Enemy voice #4
- 25) Enemy voice #5

### **Music:**

- 26) Main Theme, Major & Minor scale (multiple variations & durations)
- 27) Suspense Drumfill
- 28) Suspense Drones
- 29) Suspense Strings
- 30) Reward Fanfare

## **Stage 3**

### **Stage 3: "Oil Platform Takedown"**

- 1) Waves
- 2) Wind (brisk, swirling, continuous with occasional gusts)
- 3) Footsteps (various)
- 4) Beeping of clocked timer on bombs
- 5) Buzzing of oil generator
- 6) Explosions (various)
- 7) Gunshots (various)
- 8) Doors opening/closing
- 9) "Thump" ground impact when a soldier collapses
- 10) Heartbeat (P.O.V.)
- 11) Bullets Ricocheting off surfaces (various)
- 12) Helicopter

(continued)

# **SOCOM U.S. Navy Seals : Sound Count Sheet**

## **Stage 3: "Oil Platform Takedown"**

### **Speech:**

- 13) "HQ" voice #1 (female)
- 14) "HQ" voice #2 (male)
- 15) Kahuna (Able unit, first player character)
- 16) Boomer (Able unit)
- 17) Jester (Bravo unit)
- 18) Spectre (Bravo unit)
- 19) Enemy voice #1
- 20) Enemy voice #2
- 21) Enemy voice #3
- 22) Enemy voice #4
- 23) Enemy voice #5

### **Music:**

- 24) Main Theme, Major & Minor scale (multiple variations & durations)
- 25) Suspense Drumfill
- 26) Suspense Drones
- 27) Suspense Strings
- 28) Reward Fanfare

## **APPENDIX C**

# SSX Tricky: Sound Count Sheet

## Game Description:

*SSX Tricky is coming at you with insane, sick UBER tricks, surreal mind-blowing worlds, and a cast of funky-fresh characters. Blast down the tracks at all-out speeds and suck up the biggest adrenaline you've ever inhaled. All of your favourite SSX courses are tweaked for more speed, more elevation, and more insane thrills, plus two wild new tracks that will blow your mind. Hit the mountain harder with SSX Tricky. The sky is your stage. (© 2001 Electronic Arts Inc. SSX Tricky Game Instruction Manual. p.5)*

Catalogued below are the sounds documented during the analysis of *SSX Tricky*. The sounds were isolated for the purpose of deconstructing the sound components in the textual analysis. This list does not attempt to accurately state all of the sound samples employed by the designers, but rather seeks to illuminate the key sounds that are perceived by the player and pertinent within the overall construct of the virtual sonic gaming environment for the purposes of this analysis. The "stages" in *SSX Tricky* represent different global venues rather than sequential levels in the game, with terrain courses increasing in difficulty to challenge players as their play style improves.

## Stage 1, 2 & 3

### Stages 1-3: Tracks "Garibaldi", "Snowdream" & "Elysium Alps"

#### SFX

- 1) Crowd\*
- 2) "Beep" countdown sound
- 3) Wind (during "airtime")\*
- 4) Snowboard Glide\* (a) snow (b) ice (c) dirt (d) rock (e) tree (f) rail (g) pole
- 5) Snowboard Carve/Scrape\*(a) snow (b) ice (c) dirt (d) rock (e) tree (f) rail (g) pole
- 6) Snowboard Grind\* (a) snow (b) ice (c) dirt (d) rock (e) tree (f) rail (g) pole
- 7) Footsteps on board (when doing a trick)
- 8) Collision (board impact)\* (a) snow (b) ice (c) dirt (d) rock (e) tree (f) rail (g) pole
- 9) Collision (body impact)\* (a) snow (b) ice (c) dirt (d) rock (e) tree (f) rail (g) pole
- 10) "Swash" sound when jumping
- 11) Drone (when player gets "re-positioned")
- 12) Snowmachine
- 13) Snowmobile
- 14) Chairlift
- 15) Glass breaking
- 16) "Jingle" (when game is paused or restarted)
- 17) Bell (with each marker on "Tricky level meter")
- 18) Siren (when reaching maximum "Tricky" level)
- 19) "Twinkle" (when getting a power-up icon)
- 20) "Crackle"\* (when brushing against tree branches)
- 21) "Thud"\* (when landing a jump)
- 22) "Buzz" (when multiple combination of tricks is achieved)
- 23) "Squeal & Explosion" (when fireworks go off)

\*these sounds are variable according to the velocity/speed of the character (fast, slow), character size (large, small), board type (wide, long) and snow type (fluffy, hard-packed).



# SSX Tricky: Sound Count Sheet

(continued)

## SFX (specific to environment):

### Track 1: "Garibaldi"

- 1) Crow squawk
- 2) Birds chirping

### Track 2: "Snowdream"

- 1) Owl hoot
- 2) Seagulls

### Track 3: "Elysium Alps"

- 1) Birds chirping
- 2) Water splash (when landing in river)
- 3) Crow sound #1
- 4) Crow sound #2
- 5) Seagulls
- 6) Wolf howl

## Speech/Voices:

### Announcer Voices:

- |                        |                          |
|------------------------|--------------------------|
| 1) MC/Announcer voice: | Rozell (Rahzel) M. Brown |
| 2) Front End Voice:    | Kathleen Barr            |
| 3) Narrator Voice:     | Brent Chapman            |
| 4) Voice Direction:    | Steve Rechtshafner       |

### Character Voices:

- |             |                    |
|-------------|--------------------|
| 1) Eddie:   | David Arquette     |
| 2) Seeiah:  | Macy Gray          |
| 3) Elise:   | Lucy Liu           |
| 4) Zoe:     | Bif Naked          |
| 5) Psymon:  | Jim Rose           |
| 6) Luther:  | Oliver Platt       |
| 7) Marisol: | Patricia Valasquez |
| 8) Brodi:   | Billy Zane         |
| 9) Moby:    | Nick Malaperiman   |
| 10) Mac:    | Ryan Wall          |

©2001 Electronic Arts, Inc. *SSX Tricky* "Credits", Game Instruction Manual. P.30-31.

# *SSX Tricky: Sound Count Sheet*

(continued)

<b>Music:</b>		
<b>Artist</b>	<b>Song Title</b>	<b>Record Company</b>
1) RUN DMC	It's Tricky	Protoons/Arista/Profile
2) Shocore	Bone Cracker	Shocore Music Inc.
3) Mixmaster Mike	Board Burner	Mixmaster Mike
4) Mixmaster Mike/Rahzel	Slayboarder	MCA/Mixmaster Mike
5) BT	Smartbomb	BMI/Nettwerk
6) BT	Hip Hop Phenomenon	BMI/Nettwerk
7) Skank	Shake What Yo Mama Gave Ya Control/7	even Music
8) Space Raiders	Song For Dot	Skint/Sony/BMI
9) Aphrodite	King of the Beats	EMI/V2
10) The Forth	Reality Detached	Control/Quad
11) Rasmus	Peaktime	Big Life/Bolshoi
12) Rasmus	Superwoman	Big Life/Bolshoi
13) Huda Hudia	System Overload	Kaleidoscope
14) Hybrid Holmes	Finished Symphony	Distinctive/Sherlock
15) Bif Naked	Leader	Taykar/H.R.M./Lava
16) Black Label Society	The Rose Petalled Garden	BMI/Zakk Wylde
17) Pizzicato 5	Baby Portable Rock	Matador/Doormat
18) Chris & Martin Stieber	Twin Peak Loop	BMG/Stieber
19) John Morgan	Top Bomb	Electronic Arts
20) John Morgan	Downtime 2001Remix	Electronic Arts
21) John Morgan	Bass Invaders	Electronic Arts
22) John Morgan	Adam's Revenge	Electronic Arts
23) John Morgan	Gin and Sin	Electronic Arts
24) John Morgan	Speed Freak - Snowdream Intro	Electronic Arts
25) John Morgan	Elysium Intro	Electronic Arts
26) John Morgan	Renegade - Merquy City Intro	Electronic Arts
27) John Morgan	Himalayas - Mesablanca Intro	Electronic Arts
28) John Morgan	Destroy the Competition	Electronic Arts
29) John Morgan	Garibaldi Theme	Electronic Arts
30) John Morgan	Alaska Theme	Electronic Arts
31) John Morgan	Equinox - Megaplex Theme	Electronic Arts
32) John Morgan	Brodi Theme	Electronic Arts
33) John Morgan	Elise Theme	Electronic Arts
34) John Morgan	JP Theme	Electronic Arts
35) John Morgan	Mac Theme	Electronic Arts
36) John Morgan	Marisol Theme	Electronic Arts
37) John Morgan	Metrognome - Moby Theme	Electronic Arts

©2001 Electronic Arts, Inc. *SSX Tricky* "Credits", Game Instruction Manual. p.31-34.

## **APPENDIX D**

# MVP Baseball 2003: Sound Count Sheet

## Game Description:

*Bottom of the ninth, bases loaded, two outs, and you're at the plate... Experience the intensity of the big-league pitcher-batter duel with MVP Baseball 2003 from EA Sports™. With a totally new look and feel, MVP Baseball 2003 delivers the excitement and drama of baseball's defining moment – the pitcher-batter showdown. Featuring true-to-life baseball gameplay, MVP Baseball 2003 redefines what a baseball game should look and play like with new player animations, Franchise Mode play, game depth, and much more.*

(©2003, Electronic Arts Inc. "Introduction", MVP Baseball 2003 Game Instruction Manual, p.5)

Below is the list of sounds documented during the analysis of MVP Baseball 2003. The sounds were isolated and catalogued for the purpose of deconstructing the sound components in the textual analysis. This list does not attempt to accurately state all of the sound samples employed by the designers, but rather seeks to illuminate the key sounds that are perceived by the player and pertinent within the overall construct of the virtual sonic gaming environment for the purposes of this analysis. Sports Sim Games do not feature "stages" to vary levels of difficulty. Instead, games are played in the stadium venue appropriate to the home team of a given match, or stadiums can be manually chosen by the player.

### SFX:

- 1) Crowd\*
- 2) Crowd Voices (various)\*
- 3) Hawkers & Vendors (various)
- 4) Organ
- 5) Foghorn
- 6) "Crack" of the ball hitting the bat\*
- 7) "Whoosh" of the pitch\*
- 8) "Whiff" of the bat swinging\*
- 9) Ball in glove impact sound
- 10) Bat dropping sound
- 11) Runner sliding into base\*
- 12) "Thud" of ball impact with outfield wall
- 13) "Clunk" sound of ball hitting the batter
- 14) "Buzz" sound when game is paused/unpaused
- 15) "Boink" sound when menu screen changes

### Venue (Stadium) Specific SFX:

- 16) Crowds Chanting\*
- 17) Vendor Hawking

### Player Specific SFX:

- 18) Crowd Voices\*

\*These sounds are variable according to input parameters. They vary according to environmental characteristics, the player's style of play, and/or according to the intensity of the gamestate.

(continued)

# MVP Baseball 2003: Sound Count Sheet

## Speech/Voices:

- 1) Play-by Play Commentator: Duane Kuiper
- 2) Colour Commentator: Mike Krukow
- 3) PA Announcer Voice: Joe Richards
- 4) Tutorial Voice Direction: Harold Reynolds
- 5) Umpire 1
- 6) Umpire 2

©2003 Electronic Arts, Inc. MVP Baseball 2003 (video game) Game credits screen.

## Music:

Artist	Song Title	Record Company
1) All-American Rejects	Swing, Swing	SKG/Dreamworks
2) The Donnas	Who Invited You?	BMI/Atlantic
3) The Exies	Without	WB/BMI/Virgin
4) (hed) Planet Earth	Blackout	Zomba/Jive
5) Revis	Caught in the Rain	BMI/Epic/Sony
6) Boysetfire	Handful of Redemption	Flandrin/WindUp
7) Burning Bridges	Arctic Snow	ASCAP/V2
8) OK Go	Don't Ask Me	BMI/Capitol/EMI
9) Pacifier	Bullitproof	Dreamworks/Arista
10) Shinedown	Fly from the Inside	Control/BMI/Atlantic
11) Socialburn	Everyone	BMI/Elektra
12) Soundtrack of our Lives	Sister Surround	MCA/Universal
13) Sum 41	All Messed Up	Chrystalis/EMI/Def Jam
14) Taproot	Poem	Universal/Atlantic
15) Morning Maker	Are You Ready Now	SOCAN/MorningMakerMusic
16) Superheist	Crank the System	Shock Music
17) Just Blaze	Crunch Time	F.O.B./ASCAP/EA Inc.
18) Assisted Living	Gimme	Alcala, Rumblefish
19) Five Horse Johnson	Lightning When In Need	Small Stone/ASCAP
20) Johnny Ferreira	Mexico	SOCAN/Pair-a-Dice
21) Banditos Bonitos	La Colegiala	Planet Audio
22) MC Mario	Raise the Roof	Wild Noise/Mastermind
23) The Waking Hours	Revenge	ASCAP
24) Border Ruffians	Who's the Man	Dumb Guy/ASCAP
25) Shiftkit	Waiting	JKD/SOCAN

©2003 Electronic Arts, Inc. MVP Baseball 2003 (video game) Game Credits Screen.

# BIBLIOGRAPHY

## BOOKS:

- Adorno, Theodor; Horkheimer, Max. Dialectic of Enlightenment. Translated By John Cumming. New York: Herder and Herder, 1972.
- Arato, Andrew; Gebhardt, Eike. "On the Fetish-Character in Music and the Regression of Listening"; The Essential Frankfurt School reader (Chapter by Theodor W. Adorno). NY: Urizen Books, c1978.
- Attali, Jacques. Noise: The Political Economy of Music. The University of Minnesota Press, 1985.
- Berger, Arthur Asa. Video Games: A Popular Culture Phenomenon. Transaction Publishers, NB, London: 2001.
- Bregman, Albert. Auditory Scene Analysis : The Perceptual Organization of Sound. Cambridge, MIT Press, 1990.
- Cope, David. The Algorithmic Composer. A-R Editions, Inc. Madison, Wisconsin, 2000.
- Held, David. Introduction to Critical Theory. London: Hutchinson, 1980.
- Hurd, Michael. The Composer. New York & Toronto: Oxford University Press, 1968.
- Kent, Steven L. The Ultimate History of Video Games. Prima, Random House: New York, 2001.
- Lomax, Alan. "Song Structure and Social Structure," Ethnology, Vol. 1. Columbia University, 1962
- Marks, Aaron. The Complete Guide to Game Audio. Lawrence: CMP Books, 2001.
- Murray, Janet H. Hamlet on the Holodeck: The Future of Narrative in Cyberspace The MIT Press, Cambridge, 1997.
- Negroponte, Nicholas. Being Digital. London: Hodder & Stoughton, 1995.
- Ostwald, Peter F. Communication and Social Interaction: Clinical and Therapeutic Aspects of Human Behavior. NY, London: Grune & Stratton, c1977.
- Packer, Randall; Jordan, Ken, eds. Multimedia: From Wagner to Virtual Reality. New York : Norton, 2001.
- Pellman, Samuel. An Introduction to the Creation of Electroacoustic Music. Wadsworth Publishing, California, 1994.

- Rowe, Robert. Machine Musicianship. MIT Press, Cambridge, 2001.
- Schafer, R. Murray. The Soundscape: The Tuning of the World, Destiny Books, Vermont, 1994, 1977.
- Summers, Jodi. Interactive Music Handbook: The Definitive Guide to Internet Music Strategies, Enhanced CD Production, and Business Development. New York: Allworth Press, c1998.
- Swingewood, Alan. The Myth of Mass Culture. London : Macmillan, 1977.
- Truax, Barry. Acoustic Communication. Ablex Publishing, NJ: 2001, 2<sup>nd</sup> ed.
- Webster. Webster's Encyclopedic Unabridged Dictionary of the English Language. Dilithium Press, Portland House, NY, 1989.
- Wiener, Norbert. The Human Use of Human Beings: Cybernetics and Society. 2<sup>nd</sup> ed. Garden City, NY: Doubleday, 1954.
- Winkler, Todd. Composing Interactive Music: Techniques and Ideas Using MAX. The MIT Press, Cambridge. 1998.

#### ARTICLES & JOURNALS:

- Bloomberg, "Film-Based Games Drive Sales Up by 48 Per Cent" *The Vancouver Sun*, January 31, 2003. Section D4.
- Jones, Christopher. "Game for New Music" *Music Publisher Canada*, Fall 2002 Issue. p.4-5
- Kline, Stephen. "Killing Time? A Canadian Meditation on Video Game Culture," in *Children and Media Violence*. UNESCO Clearinghouse Annual, 2000.
- McCullough, Michael. Technology: "Entertainment goes Wireless", *Vancouver Sun*, Saturday, March 22, 2003. Section E1.
- Truax, Barry. "Musical Creativity and Complexity at the Threshold of the 21<sup>st</sup> Century" from Interface, Vol. 21, 1992. p.29-42.
- Truax, Barry. "The Aesthetics of Computer Music: A Questionable Concept Reconsidered." Organised Sound, 2000. p.119-126.
- Westerkamp, Hildegard. "Listening and Soundmaking: A Study of Music-as-Environment" from Sound By Artists Eds. Lander & Lexier, published by Art Metropole and Walter Phillips Gallery, Alberta, Canada: 1990.

## WEB REFERENCES:

- Ahl, David H. "Editorial." 1983. <http://cvmm.vg-network.com/vag1.htm>. Accessed: April 26, 2003.
- Andresen, Gavin. "Playing by Ear: Creating Blind-Accessible Games." May 20, 2002. [www.gamasutra.com/resource\\_guide/20020520/andresen\\_01.htm](http://www.gamasutra.com/resource_guide/20020520/andresen_01.htm).
- Arnold, Jim. (Business Wire: Dolby Press Release) "Dolby Digital 5.1 Turns Up the Volume..." San Francisco, August 20, 2002. [http://biz.yahoo.com/bw/020820/200266\\_1.html](http://biz.yahoo.com/bw/020820/200266_1.html).
- Belinkie, Matthew. "Video game music: not just kid stuff." *Video Game Music*. December 15, 1999. <http://www.vgmusic.com/vgpaper.shtml>.
- Bernstein, Daniel. "Creating an Interactive Audio Environment." Published in "Game Developer Magazine", October 1, 1997. [http://www.gamasutra.com/features/sound\\_and\\_music/111497/audio\\_tips\\_01.htm](http://www.gamasutra.com/features/sound_and_music/111497/audio_tips_01.htm)
- BlueGamers: <http://www.bluegamers.com/ps2/reviews/mohf>. Accessed: May 22, 2003.
- Boecker, Thomas. Merregnon Studios: "What is Merregnon?" <http://www.merregnon.com/main.php>. Accessed: June 19, 2003.
- Brandon, Alexander. "Interactive Music: Merging Quality with Effectiveness." March 27, 1998. Vol.2, Issue 13. [http://www.gamasutra.com/features/sound\\_and\\_music/19980327/interactive\\_music\\_quality\\_intro.htm](http://www.gamasutra.com/features/sound_and_music/19980327/interactive_music_quality_intro.htm)
- Clark, Andrew. "An Interview with Darryl Duncan". *Gamasutra*, May 15, 2001. [http://www.gamasutra.com/resource\\_guide/20010515/marks\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/marks_01.htm).
- Clark, Andrew. "Designing Interactive Audio Content To-Picture." *Gamasutra*, December 20, 1999. [http://www.gamasutra.com/features/19991220/clark\\_01.htm](http://www.gamasutra.com/features/19991220/clark_01.htm).
- Coleman, Christopher. "A Conversation with Composer Michael Giacchino." 2001. <http://www.tracksounds.com/specialfeatures/Interviews/interviewing.htm>. Accessed: October 17, 2002.
- DENON Ltd., 2003. FAQ Website: *DTS* <http://denon.jp/asia/faq/2-8.html>. Accessed: April 30, 2003.
- Electronic Arts Inc.; "Investor Info." [http://www.corporate-ir.net/ireye/ir\\_site.zhtml?ticker=ERTS&script=2100](http://www.corporate-ir.net/ireye/ir_site.zhtml?ticker=ERTS&script=2100). Accessed: May 16, 2003.
- Electronic Arts Inc.; "About EA." 2002 <http://www.ea.com/global/corporate/aboutea.jsp?src=11001homecorpabounone>. Accessed: May 16, 2003.



- Gamasutra.com, "Features, Audio & Graphics" (various 1997-2003):  
[http://www.gamasutra.com/php-bin/article\\_display.php?category=view\\_all](http://www.gamasutra.com/php-bin/article_display.php?category=view_all). Accessed: September, 2002.
- Game Development Search Engine (GDSE):  
<http://www.gdse.com/servlet/gdse.links?s=65>. Accessed: August 18, 2002.
- GameDevelopers.net: *Music & Sound Resource*.  
<http://www.gamedev.net/reference/list.asp?categoryid=51>. Accessed 18, 2002:
- Gamespot.com: "Gamespot Presents: Xbox Dossier." 2002.  
[http://gamespot.com/gamespot/features/video/x\\_box\\_dossier/p1\\_01.html](http://gamespot.com/gamespot/features/video/x_box_dossier/p1_01.html). Accessed: May 12, 2003.
- Gamespy, 2002: <http://www.gamespy.com/SOCOM/inside/gamefeatures.asp>. Accessed: May 12, 2003.
- Golberg, Marty. "The History of Computer Gaming Part 1: Forgotten Origins."  
<http://www.classicgaming.com/features/articles/computergaminghistory>. Accessed: January 29, 2003.
- Harland, Kurt. "Composing for Interactive Music." Feb 17, 2000.  
[http://www.gamasutra.com/features/20000217/harland\\_01.htm](http://www.gamasutra.com/features/20000217/harland_01.htm).
- Hays, Tom. "DirectMusic for the Masses." Gamasutra, November 6, 1998. Vol.2, Issue 44.  
[http://www.gamasutra.com/features/sound\\_and\\_music/19981106/directmusic\\_01.htm](http://www.gamasutra.com/features/sound_and_music/19981106/directmusic_01.htm).
- Hill, Greg. "Capturing Engine Sounds for Games." Gamasutra, October 30, 2002.  
[http://www.gamasutra.com/features/20021030/hill\\_01.htm](http://www.gamasutra.com/features/20021030/hill_01.htm).
- Hyde-Smith, Alex. "House eJay" *Music4games.net*. October 16, 2002.  
[http://www.music4games.net/f\\_ejay\\_house.html](http://www.music4games.net/f_ejay_house.html).
- IGN Insider: "Sound" *MVP Baseball 2003 review*, IGN.com, March 11, 2003.  
<http://ps2.ign.com/articles/388/388870p4.html>.
- Labrecque, Denis. "Altering Algorithms to Create '3D' Sound." EE Times, December 11, 2001.  
<http://www.etimes.com/story/OEG20011218S0047>.
- Lagim, Bjorn Arve. "The Music of Anarchy Online: Creating Music for MMOGs."  
 Gamasutra.com, 2002.  
[http://www.gamasutra.com/resource\\_guide/20020916/lagim\\_01.htm](http://www.gamasutra.com/resource_guide/20020916/lagim_01.htm). Accessed: May 12, 2003.
- Lewis, George E. "Singing the Alternative Interactivity Blues." Western Front magazine, and the Association for the Advancement of Creative Musicians website: 1995  
<http://audiolabo.free.fr/revue1999/content/lewis2.htm>. Accessed: January 16, 2003.
- Macy, Laura (ed.) "Film Music." *Grove Music*. 2<sup>nd</sup> ed. The New Grove Dictionary of Music and Musicians. May 12, 2002 <http://www.grovemusic.com/index.html>.

- Marks, Aaron. "An Interview with Will Davis." Gamasutra.com.  
[http://www.gamasutra.com/features/20000901/Davis\\_01.htm](http://www.gamasutra.com/features/20000901/Davis_01.htm).  
Accessed: January 21, 2003.
- Marr, John. "The Game AI Page." Game AI.com, 2003. <http://www.gameai.com/ai.html>.  
Accessed: March 27, 2003.
- McCarthy, Ryan. "SOCOM: U.S. Navy Seals PS2 Staff Review", 2002.  
<http://www.psx2central.com/reviews/psx2socom.htm>. Accessed: March 27, 2003.
- McDonald, Glenn. "Gamespot: Brief Timeline of Game Music."  
[http://gamespot.com/gamespot/features/video/vg\\_music/index.html](http://gamespot.com/gamespot/features/video/vg_music/index.html). Accessed:  
February 4, 2003.
- Miller, Mark Steven. "Producing Interactive Audio" Game Developer Magazine,  
October 13, 1997.  
[http://www.gamasutra.com/features/sound\\_and\\_music/111497/interaudio1.htm](http://www.gamasutra.com/features/sound_and_music/111497/interaudio1.htm).
- Minnicino, Micheal J. The New Dark Age: The Frankfurt School and "Political Correctness." Courtesy of the *Schiller Institute*. From the Winter, 1992 issue of *FIDELIO Magazine*. [http://www.schillerinstitute.org/fid\\_91-96/921\\_frankfurt.html](http://www.schillerinstitute.org/fid_91-96/921_frankfurt.html).
- Music4games.net "EA Trax Interview: Steve Schnur" October 15, 2002.  
[http://www.music4games.net/f\\_eatrax.html](http://www.music4games.net/f_eatrax.html).
- O'Donnell, Marty. "Producing Audio for Halo." Gamasutra.com, October 30, 2002.  
[http://www.gamasutra.com/resource\\_guide/20020520/odonnell\\_01](http://www.gamasutra.com/resource_guide/20020520/odonnell_01).
- O'Neill, Cliff. "SSX Tricky Game Review." Sports Gaming Network.  
[http://www.sgn.cc/other/ssx\\_tricky\\_ps2/review\\_ps2.shtml](http://www.sgn.cc/other/ssx_tricky_ps2/review_ps2.shtml). Accessed: May 11, 2003.
- Patterson, Scott. "Interactive Music Sequencer Design" Gamasutra, May 15, 2001.  
[www.gamasutra.com/resource\\_guide/20010515/patterson\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/patterson_01.htm).
- Payne, Joseph. "*FilmTracks* Modern Soundtrack Reviews: Medal of Honor."  
[www.filmtracks.com/titles/medal\\_honor.html](http://www.filmtracks.com/titles/medal_honor.html). Accessed: November 18, 2002.
- Pellis, Jay. "The Console Gaming Phenomenon."  
<http://www.geocities.com/sbobbitt21/audyssey24/console.html>. Accessed: March 27,  
2003.
- Pidkameny, Eric. "Levels of Sound" *Video Game Music* archive. May 15, 2002.  
[www.vgmusic.com/information/vgpaper2.html](http://www.vgmusic.com/information/vgpaper2.html).
- Playstation.com. Sony Computer Entertainment America Inc.  
<http://www.us.playstation.com/games/SCUS-97134.asp>. Accessed: April 22, 2003.

Publius, "Norbert Wiener Quotes" (taken from The Human Use of Human Beings: Cybernetics and Society. 2<sup>nd</sup> ed., 1954) <http://www.angelfire.com/co/1x137/wiener.html>. Accessed: April 12, 2003.

Research and Development of the US Dept of Energy, 1981.  
<http://www.osti.gov/accomplishments/videogame.html#pagetop>.  
Accessed: February 4, 2003.

Ross, Rob. "Interactive Music...er, Audio" Gamasutra, May 15, 2001.  
[http://www.gamasutra.com/resource\\_guide/20010515/ross\\_01.htm](http://www.gamasutra.com/resource_guide/20010515/ross_01.htm).

Sanger, George Alistair. "The Sound of Money (Down the Potty): Common Audio Mistakes in Kids' Games." Gamasutra, May 15, 2001.  
[http://www.gamasutra.com/features/20010515/sanger\\_01.htm](http://www.gamasutra.com/features/20010515/sanger_01.htm).

Satterfield, Shane. "Gamespot Presents: GameCube Dossier". *Gamespot*. May 10, 2002.  
[http://gamespot.com/gamespot/features/video/gamecube\\_dossier/p2.html](http://gamespot.com/gamespot/features/video/gamecube_dossier/p2.html).

Special Interest Group Graphics (SIGGRAPH): <http://www.siggraph.org/gen-info>. Accessed: February 7, 2003.

Thomas, Chance. "Day One Audio" Published in *music4games.net*, October 16, 2002.  
[http://www.music4games.net/f\\_dayoneaudio.html](http://www.music4games.net/f_dayoneaudio.html).

Thomas Jr., Donald A. "1980s" *PC & Game History*. August 08, 2002  
<http://www.icwhen.com>.

Warner Bros. Inc. "Enter the Matrix: Game Info."  
<http://www.enterthematrix.com/html/index/html>. Accessed: June 2, 2003.

Winter, David. "Pong Story" <http://www.pong-story.com/rhbaer.htm>. Accessed: February 16, 2003.

Yackley, David. "Microsoft DirectMusic: Creating New Musical Possibilities." Microsoft Corporation, December 22, 1999.  
[http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnmusic/html/dm\\_nmp.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnmusic/html/dm_nmp.asp).

### **MULTIMEDIA APPLICATIONS:**

Cope, David. The Algorithmic Composer. CD-ROM. A-R Editions, Inc. Madison, Wisconsin, 2000.

Electronic Arts, Inc. *MVP Baseball 2003*. (video game) ©2003 EA Sports™, Electronic Arts, Inc.

Electronic Arts, Inc. *SSX Tricky*. (video game) ©2001 EA BIGSports™, Electronic Arts, Inc.

Sony Computer Entertainment America Inc. *SOCOM: U.S. Navy SEALS*. (video game) ©2002 Sony Computer Entertainment Inc.

Microsoft Corporation, "Sound Recording and Reproduction." *Microsoft® Encarta® Encyclopedia*, 2001.

Sony Computer Entertainment Inc. "Behind the Scenes, SOCOM: U.S. Navy Seals": *PlayStation® Underground Jampack*. © 2002 SCEAI.

Truax, Barry. "Handbook for Acoustic Ecology" CD-ROM, Cambridge Street Publishing, 1999.

Winkler, Todd. Composing Interactive Music: Techniques and Ideas Using MAX. CD-ROM. The MIT Press, Cambridge. 1998.