

## An Introduction to the Participatory and Non-Linear Aspects of Video Games Audio<sup>1</sup>

Karen Collins (www.gamessound.com)

It has been estimated that the average Westerner hears music for three and a half hours per day, most of which is in a linear form.<sup>2</sup> Like being locked onto a straight train track, it has been composed to start at one point and progress to another point.<sup>3</sup> A composer of music for linear media can predict how the music will sound from beginning to end for the listener, and compositions are constructed with this aspect taken for granted. The music of non-linear media like video games, however, works more like a major urban metro: At any time, we may want to be able to hop off at one station and hop onto another train going in a new direction. We may not get on at the end of this new train, but perhaps on one of the middle cars. The train may choose to speed up at night, or slow down through built-up urban areas. Every audio cue (train car) must be designed to stand alone, since there is no way to predict its hundreds of possible directions: There is no “correct” sequence of events for the train to follow. A unique relationship arises, then, between these cars and tracks, working with and connecting to one another.

The non-linear aspects of games audio, along with the different relationship the audio has with its audience, poses interesting theoretical problems and issues. Being an audio-visual form consumed on screen, it may seem at first logical in researching games audio to draw upon film and television theory; however, there are very distinct differences between these media for which new terminology and new theoretical approaches must be considered.<sup>4</sup> Although there has been significant academic research into related areas of multimedia and audio in terms of technology, communication, and development, work into the sonic aspects of audio-visual media has neglected games. Similarly, studies and theories of video games have, for the most part, disregarded the audio.<sup>5</sup> While there have been a scattering of articles published sporadically in the last few years, video games audio remains largely unexplored.<sup>6</sup> It is the aim of this paper, then, to begin to lay a few theoretical foundation stones upon which further research can be built, and to introduce to the reader some of the implications of the participatory nature of gaming in terms of both practical and theoretical perspectives.

To illustrate the discussion, I focus on two games, as they have both enjoyed considerable popularity and are easily obtainable to any researcher who may wish to examine them.<sup>7</sup> The first game, *Grim Fandango*, is an award-winning film-noir LucasArts PC adventure game released in 1998 and set in a Mardi Gras atmosphere in the land of the dead. The main character, Manny Calavera, is a travel agent for the Department of Death (D.O.D) who sets out on a journey to expose the corruption in the corporation, with the help of a rebel group known as the Lost Souls’ Alliance. The game was designed by Tim Schafer, with music composed by Peter McConnell.<sup>8</sup> The second game is *The Legend of Zelda: Ocarina of Time* (hereafter discussed as *Zelda*), released by Nintendo in 1998 to much acclaim, selling almost nine million copies, and eventually achieving the number two spot on IGN Entertainment’s ‘Top 100 Games of All Time’.<sup>9</sup> It was later re-released with extended sequences as a bonus disk with *The Legend of Zelda: The Wind Waker* (Nintendo 2003). *Zelda* was so popular that instruments “inspired by” or “designed after” the sweet potato ocarina have since been sold, and one store even includes songs from the game in their marketing.<sup>10</sup> Created by Shigeru Miyamoto and composed by Koji Kondo, with sound design by Takuya Maekawa and Yoji Inagaki, the game follows the lead character, Link, through an adventure to save the land of Hyrule from Ganondorf, the king of evil who has invaded the Sacred Realm and stolen the Triforce of Power. Both of the games used to illustrate this paper are from the adventure genre, which requires puzzle solving skills and patience. These types of games have a tendency to be drawn out and complex, and the audio in the games is critical in helping the player to adapt to the many challenges that await them. Although audio in the adventure genre tends to be more elaborate than, for instance, simulators or sports games, the same functions and characteristics discussed below will generally apply. Nevertheless, it must be noted that audio can play a more limited role in certain genres of games, as I will discuss more fully below.

Video games sound is often referred to in a vague manner as “interactive”, “adaptive” or “dynamic”. An introductory text to Microsoft’s audio development tool DirectX 9 suggests that these terms have different meanings, as it promises “software engineers and content creators the ability to create interactive, adaptive, *and* dynamic audio content” (Fay et al. 2004, back cover; my emphasis). A few attempts by those in the industry have been made to narrow down the meaning of the terms to refer to specific ways in which audio relates to a user in the context of gaming (see Bajakian et al 2003; Fay et al. 2004; Whitmore 2003). Although these more narrowly defined terms are not yet widely used, it is worth making the distinctions here.

*Interactive* audio refers to sound events occurring in reaction to gameplay, which can respond to the player directly. In other words, if for instance a player presses a button, the character on screen swings their sword and

makes a “swooshing” noise. Pressing the button again will cause a recurrence of this sound. The “swoosh” is an interactive sound effect. *Adaptive* audio, on the other hand, “reacts appropriately to—and even anticipates—gameplay” rather than responding directly to the user (Whitmore 2003). As Todd Fay indicates, “in many ways, it is like interactive audio in that it responds to a particular event. The difference is that instead of responding to feedback from the listener/player, the audio changes according to changes occurring within the game or playback environment” (Fay et al. 2004, 6). An example is *Super Mario Brothers* (Nintendo 1985), where the music plays at a steady tempo until the time begins to run out, at which point the tempo doubles. I use the term *dynamic* audio to encompass both interactive and adaptive audio. Dynamic audio, then, is, audio which reacts to changes in the gameplay environment *or* in response to a user. There are several different ways in which dynamic audio interacts with a listener/player, and in which it acts within a game’s diegesis, as outlined below.

### Degrees of Dynamic Activity in Games Audio

Dynamic audio complicates the traditional diegetic/non-diegetic division of film sound.<sup>11</sup> The unique relationship in games posed by the fact that the audience is engaging directly in the sound-making process onscreen (discussed more below), requires a new type of categorisation of the image-sound relationship. Games sound can be categorised broadly as diegetic or non-diegetic,<sup>12</sup> but within these broad categories can be separated further into non-dynamic and dynamic sound, and then divided further still into the types of dynamic activity as they relate to the diegesis and to the player.

Apart from cut-scenes (movie sequences in which the player’s input functions are impeded while a short clip to advance the plot is played), which are fixed in a linear fashion, the degrees of dynamic activity in a game are sometimes fluid, posing further difficulty in the classification of the sounds. For instance, in *Zelda’s* Kokiri Forest, during the first portion of the game, we are continuously in daytime mode as we get trained in gameplay, and the Kokiri theme that plays throughout does not change except at those points where a player enters a building or encounters an enemy. While interactive, it is not adaptive at this point. After we complete our first major task and arrive at the next portion of the game (there are no distinct “levels” in this game), we then experience the passing of time, and can return to the forest. Now, if we return at night, the music has faded out to silence. At dawn, it will return to the main theme: the theme has become adaptive. In other words, a cue which is interactive or adaptive at one point in the game does not necessarily remain so throughout. Similarly, in *Asheron’s Call 2: The Fallen Kings* (Turbine Software 2003), the non-diegetic music that plays in the background of scenes becomes diegetic when players decide to have their character play an instrument or sing along with the music. Not only has the music changed from non-dynamic to interactive, but it has also gone from non-diegetic to diegetic. As such, then, although I have distinguished levels of sound here, they must be viewed as fluid, rather than fixed, for many types of audio cues.

The most basic level of non-diegetic audio for games is the non-dynamic linear sounds and music found most frequently in the introductory movies or cut-scenes. In these cases, the player has no control over the possibility of interrupting the music (short of resetting or turning off the game).<sup>13</sup> In the introduction to *Zelda*, for instance, a short a dream sequence movie is played, explaining the plot. If the player does not start the game (leading to further cut-scenes), the entire introduction sequence loops. Similar plot advancement movies are spliced into *Grim Fandango*. At key points in the game, a pre-set cut-scene movie loads, leading us to the next stage in the plot. For instance, Manny meets with Salvador, the revolutionary, in his underground hideout to conspire to expose the inequities of the D.O.D. When Manny gives Salvador the moulded impression of his teeth (necessary for access to the building), a cut-scene ends that stage of the game (El Marrow) and leads us to the next location (the Petrified Forest). The music during this intermission cut-scene begins with the theme for the hideout, and then changes to that of the new location without the player’s input: It is, in other words, linear, *non-dynamic*, *non-diegetic music*.

Non-diegetic audio can also contain various levels of dynamic activity. *Adaptive non-diegetic* sounds are sound events occurring in reaction to gameplay, but which are unaffected by the player’s direct movements, and are outside the diegesis. As discussed above, the music in *Zelda* fades out at dusk and stops altogether during the night. At dawn, a quick “dawn theme” is played, followed by a return to the area’s main theme music. The player cannot re-trigger these events (except by waiting for another day to pass). *Interactive non-diegetic* sounds, in contrast, are sound events occurring in reaction to gameplay, which can react to the player directly, but which are also outside of the diegesis. In *Zelda*, the music changes in reaction to the player approaching an enemy. If the player backs off, the music returns to the original cue. If the player manages to find the trigger point in the game,

it is possible to hear both cues at the same time in the midst of a cross-fade. The player, then, controls the event cue, and can repeatedly trigger the cue, by, in this case, running back and forth over the trigger area.

There are also diegetic sounds (“source music” or “real sounds”) in games, which can be non-dynamic, adaptive, or interactive. In *non-dynamic diegetic audio*, the sound event occurs in the character’s space, but with which the character has no direct participation. These sounds of course occur in cut-scenes, but also take place in gameplay. For instance, in the underground hideout in *Grim Fandango*, Eva (a member of the resistance) is fiddling with a radio trying to tune in a particular station. Manny (the player’s character) has no contact with the radio: Its sound is diegetic, but non-dynamic. Diegetic sounds can also be adaptive and interactive. To return to the night/day division of time in *Zelda*, at dawn we hear a rooster crow, and in the “day” sequences of Hyrule Field, we hear pleasant bird sounds. When the game’s timer changes to night-time, we hear a wolf howl, crickets chirp, and various crows cawing. These sounds are diegetic and adaptive. On the other hand, *interactive diegetic sounds* occur in the character’s space, with which the player’s character can directly interact. The player instigates the audio cue, but does not necessarily affect the sound of the event once the cue is triggered. In *Grim Fandango*, there is a scene in the Calavera Café in which grease-monkey Glottis is playing a piano in the bar. If the player gives Glottis a VIP pass to the local racetracks, Glottis leaves the piano open. If the player then chooses, the main character Manny can sit down on the piano and play, triggering a pre-selected cue. More commonly, interactive diegetic sounds are sound effects, for instance, the sound Link’s sword makes when cutting, or the footsteps of characters in games.

Finally, a level of even more direct audio interaction is that of *kinetic gestural interaction* in both diegetic and non-diegetic sound, in which the *player* (as well as the character, typically), bodily participates with the sound on screen. At its simplest level, a joystick or controller could be argued to be kinetically interactive in the sense that a player can, for instance, play an ocarina by selecting notes through pushing buttons on a controller; but more significantly, here I refer to when a player may physically, gesturally mimic the action of a character, dancer, musician, etc. in order to trigger the sound event. In other words, the player must physically play a drum in *Donkey Konga* (Namco 2003), or play a guitar in *Guitar Hero* (Red Octane 2005), for instance. These types of games have typically required the purchase of additional equipment to play outside the traditional joystick/controller that is included with the game’s platform, although this will change with the release of Nintendo’s Wii controller in 2006, which will make kinetic gestural interaction with sound much more common. With the Wii controller, in the latest *Zelda* game, *The Legend of Zelda: The Twilight Princess* (Nintendo 2006), the player must literally swing the controller to elicit a sword movement in the game, resulting in the sword swooshing sound.

There are, as shown, several different ways in which the player is connected to, participates in, or interacts with the sound of a game, and several ways in which this sound acts internally or externally to the game’s diegesis. As the player is no longer a passive listener, but may be involved in invoking sound and music in the game, we must re-evaluate theories of reception. The notion that audiences might construct meaning from texts has come to be a subject of growing interest in cultural and media studies throughout the last few decades. The previous structuralist assumption that texts had a “preferred reading” to be decoded has given way to a post-structuralist theory of texts as having many meanings largely determined by the audience. This change in focus from the text to the audience has been referred to as the “return of the reader” or the “death of the author” and has been well documented elsewhere.<sup>14</sup> For example, if, in a film, the director wants an off-screen dog barking to inform the audience that there is a dog nearby, the sound designer records a dog bark and inserts it into the film’s mix. The audience then hears the dog bark and recognizes that there is a dog nearby in the off-screen activity. Nevertheless, the audience brings their own meanings to that sound, which may be “my dog recently died and I am sad”. In other words, the meaning is enriched by the *connotation* of the connotations. In terms of semiotic theory, I have elsewhere called this secondary level of signification “supplementary connotation” (Collins 2002). These are the unpredictable, individual and often personal connotations associated with a text.

However, even this approach does not account for media in which the audience plays an active role in the construction of the text. The traditional semiotic chain of communication from transmitter (the composer/ sound designer) to channel (the sounds) to receiver (the audience) (Figure One) is disturbed in games by the interplay between the player and the audio. In some cases, the player becomes a co-transmitter, and therefore, just as the audio in games is non-linear, it may be worth considering the communication chain as also non-linear, perhaps in a more circular fashion in which the receiver plays a more active role (Figure Two). Using the example of the dog barking, in this case, let us say that the player is in a driving game, and happens to take a curve too quickly just after the dog barks. The sound of the tires squealing is added to the transmission. In this case, the audience

may interpret that supplementary connotation one level further, as “my dog recently died when he was hit by a car and I am sad and I hate cars”, moving the train of thought away from the dog and to the car. The participatory nature of video games potentially leads to the creation of additional or entirely new meanings other than those originally intended by the creators by not only changing the reception, but also changing the transmission. We might refer to these meanings as “participatory supplementary connotations”, as the original meaning (there is a dog somewhere) is maintained, but through our own experiences and through participation is supplemented by additional meanings.

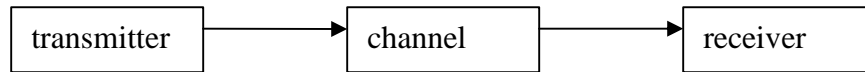


Figure One: Traditional semiotic approach to communication<sup>15</sup>

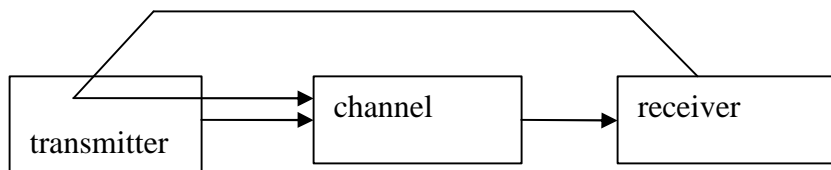


Figure Two: The impact of participation and non-linearity in gaming on communication: participatory supplementary connotations

Musical sound can, of course, also be affected by participation, as can be seen in *SingStar*, a PlayStation 2 game published by Sony in 2004, in which players can sing along with a selection of popular songs—rather similar to karaoke, although the game component comes in to play when the game scores the player based on the accuracy of their vocals to the original recording. A player may choose to have fun by not attempting a high score, but by intentionally singing off-key or in an unusual way, changing the meaning of the original song. In fact, the competitive aspect and the rhetoric of superstardom is perhaps enough to change the meanings of the song for the audience. In less obvious ways, in most games the player has an active role in the overall sound of the game. Many player actions have sonic reactions, such as the firing of a gun, jumping, etc. and while the game may dictate where these sound events occur in some places, in other places the player may simply enjoy the sound and add a few jumps or gunshots into the mix. In some cases, a game’s score is too closely synchronized to the action (a process known as “mickey-mousing” in film),<sup>16</sup> and the composer risks creating a musical “mini game” within the game. For instance, if the music’s pitch ascends as a player climbs a flight of stairs, the player may choose to run up and down the stairs to play with the music, rather than with the game (Seflon 2006). Players, then, may choose to create new, unpredictable meanings from sounds in games, and thus are part of the creation of the whole resultant sonic landscape of a game. There are clearly many theoretical implications of games sound, and I will return to this idea later. First, however, it is necessary to understand some of the practical implications of games audio for sound designers and composers.

### Dynamic Audio: Issues in Non-Linearity as it Applies to Games Sound Design and Composition

Returning now to our metaphor of the urban metro as an illustration of non-linear music’s structure, it is evident that this unpredictability may affect the audio in different ways. A player may at any time jump train cars, stop at stations, get on new trains, or change the direction of the car they are in. For composers, this unpredictable element can be very difficult to score. While scoring cues for specific gameplay events is fairly straight-forward, predicting how these cues may have to connect with other cues—what happens when two cars or tracks meet—can be very difficult. Moving smoothly from one music cue to another assists in the continuity of a game, and the illusions of gameplay, since a disjointed score generally leads to a disjointed playing experience, and the game may lose some of its immersive quality. There have been several approaches to the problems of transitioning between cues of music, and many scores use a variety of different transition types. Early games tended towards direct splicing and abrupt cutting between cues, though this can feel very jarring for the player (this was especially common in 8 and 16-bit games).<sup>17</sup> The most common transition is to fade out quickly, and then begin the new cue, or to cross-fade the two cues, which is to say, fade one out as the other cue fades in. Nevertheless,

the transition can still feel abrupt, depending on the cues or the speed of the cross-fade. *Zelda* makes frequent use of this type of segue: In places where Link enters a building, the Kokiri theme music fades out completely, and then the “house” or “shop” music cue will begin. However, as mentioned above, when Link encounters a threatening enemy, a more subtle cross-fade will occur, and it is possible to situate the character half-way between both cues at the trigger point, in order to hear both cues simultaneously. To further assist in the smoothness of the cross-fades, the *Zelda* cues drop elements as the cue fades in or out, such as the disappearance of the snare-drum in advance of other instruments in the above situation. Another common type of transition is to use a quick stinger, also known as a stab (a quick shock chord). Particularly as one of the most abrupt transitions is to that of violent combat, the stinger—combined with the sound effects of violence (gunfire, sword clashes, etc.)—will obscure any disjointed musical effects that may occur. A stinger transition is used on the roof in *Grim Fandango*, for instance, when Manny scares away the last of the pigeons. We hear a squawking trumpet blast and then the music cue changes.

There are also a few more recent attempts at transitions that are more effective, but which are far more demanding on the composer, who must think of the music in non-linear and untraditional ways. For instance, some games use cue-to-cue transitions, so that when a new cue is requested, the current sequence plays until the next marker point (perhaps the next measure, or the next major downbeat), before triggering the new cue. This type of transition can be seen in *Grim Fandango*, which uses a LucasArts patented software engine called iMuse (Interactive Music Streaming Engine). The patent describes the cue-to-cue idea as follows:

The decision points within the database comprise a composing decision tree, with the decision points marking places where branches in the performance of the musical sequences may occur... More specifically, each time a hook (allowing a jump to a new part of the sequence) or marker message is encountered in the musical sequence being played, it is compared with its corresponding value. If there is no match the message is ignored. If there is a match, a prespecified musical action occurs. (Land & McConnell 1994, 1, 14.)

The major difficulty with the cue-to-cue type of transition in most games, however, is the time lag: If the player’s character is suddenly attacked by an adversary in a game, it may be a few seconds before the music triggers this event—far too much of a delay. Nevertheless, this approach is increasingly common, as composers write literally hundreds of cue fragments for a game, to reduce transition time and to enhance flexibility in their music.<sup>18</sup> Cue-to-cue transitioning is somewhat similar to the “transition matrix” method of scoring for a game. A transition matrix contains a series of small files which enable the game engine to analyse two cues and select an appropriate pre-composed transition. The cues must use markers to indicate possible places in the score where a transition may take place. This means that the composer must map out and anticipate all possible interactions between sequences at any marker point, and compose compatible transitions; an incredibly time consuming process (Fay et al. 2004, 406).

Further methods of transitions include layering approaches to song construction, in which music is composed in instrument layers with the understanding that at any time various instruments may be dropped and others may be added. This works well with some transitions, but again, it can be difficult to move quickly from cue to cue in a dramatic change. Instrument layers are also included as part of the iMuse patent, originally intended for MIDI, although it functions in *Grim Fandango* using pre-recorded sequences of WAV files. In one example, Manny stands on the docks and the underscore plays the ‘Limbo’ cue. If the player has Manny look at the moon, he will recite a poem, and high sustained strings are added to the original cue.

An example of a game using even more complex cue-to-cue and instrument layer transitions than those found in *Grim Fandango* or *Zelda* is *Russian Squares* (Microsoft 2002), in which the goal is to clear rows of blocks by matching colours. As each row is cleared, the music responds by adding or subtracting various layers of instrument sounds. The composer, Todd M. Fay, describes his approach as follows:

*Russian Squares* uses instrument level variation to keep the individual cells from getting monotonous. Each music cell is anywhere from two to eight measures in length, and repeats as the player works on a given row... When combined with other instruments, these variations increase the amount of variation logarithmically. Most often, one to three instruments per cell use variation, and that gives the music an organic spontaneous feel, and prevents that all too familiar loopy feeling. Too much variation, however, can unglue the music’s cohesion. (Fay et al. 2004, 373).<sup>19</sup>

I have transcribed ten cues (what Fay refers to as cells) to show how the layering approach functions in *Russian Squares* (see Figure Three).<sup>20</sup> Each block represents one bar of music. The dotted lines indicate a change to a new cue sequence, which occurs with every row of blocks eliminated in the game. Cues have up to six instrument layers occurring at any time, although which layers occur in any cue is not necessarily relative to the one that comes before or after: the only constant is the steady synthesized electric bass line (instrument layer #1), though one percussion pattern is fairly consistent throughout (instrument layer #2). I have not included the addition of sound effects in the game, which occur with row clearances, time warnings, and unmoveable blocks.

Figure Three: *Russian Squares*, 'Gravity Ride': Ten cues (two minutes)

Transitions are not the only difficulty facing a games composer: non-linearity in games has many consequences which affect how a game will sound. For instance, the relationship between the different types of audio occurring in a game leads to other complications for scoring. Stockburger (2005) identifies five different categories of sound events in games: speech (dialogue), zone (ambience), score (music), effects (diegetic game sounds) and "interface sound objects" (nearly exclusive to menu screens, sounds generally "not perceived as belonging to the diegetic part of the game environment"). Particularly since dialogue and the sound effects of, for instance, combat, are all mid-range, there is a constant risk of creating a "muddy" sound if the music also contains a lot of mid-range activity. For instance, if a player's character is in an open field, talking to a non-playing character, and is supposed to hear a gunshot in the distance, if the composer has chosen to include a lot of mid-range activity in that cue (such as, for instance, snare drums and guitar), the gunshot, the conversation and/or the music is going to be obscured. Whereas this has always been problematic in film sound and other linear media, in games, the added unpredictability of where these events may occur makes the mixing far more difficult a task. There is no real-time mixing of games audio (yet): There is nobody at the mixing desk to say when the effects should drop, or the music reduced, and so the result is sometimes a clash of sounds, dialogue, music and ambience. The consequence, therefore, is that some types of games—or some particular areas within a game—must be scored with potential mixing problems in mind, and action music cues may have high and low range frequencies, but with little in the middle range. Taken out of context, then, the music can at times sound awkward or unfinished. Another equally important issue relating to this mixing problem was raised by Anahid Kassabian in her discussion of the "evaporating segregation of sound, noise and music" (2003, 92). Kassabian suggests that the soundscape of games has impacted on recent film scores, such as *The Matrix* (1999), where the distinction between sound design and underscore is greatly reduced. Similar disintegration between sound effects and music are occurring in hip-hop and electronic music where sound effects (or "non-musical sound") regularly make appearances in songs (see Collins 2002). This raises the further possibility that games sound is influencing approaches to musical production in popular and cinematic music and that the consequences—or results—of non-linearity in music extend beyond non-linear media.<sup>21</sup>

Multi-player games (games played at home with other players, or games that can be played online with other users) pose another challenge to composers. If a game is designed to change cues when a player's health score reaches a certain critical level, what happens when there are two players, and one has full health while the other is critical (Seflon 2006)? Or, if a particular sonic trigger is supposed to occur at a specific point in the game, a decision must be made as to whether it occurs when both players have entered the target area, and triggered the sound event, or when just one player has entered. There are also some interesting musical opportunities opened up by multi-player games, such as spontaneous "jam sessions" between characters/players, in a game like *Asheron's Call 2*, mentioned above, in which different species of characters play different types of instruments, and the interaction between players causes changes to the music (see Fay et al. 2004, 473-499). In this way, players are encouraged to interact with each other, and to play music together.

The length of gameplay is another critical aspect that distinguishes games from linear media, and as such also introduces new difficulties for composers. The non-linear nature of games means that gameplay length is indeterminate. A player may get stuck and never complete a game, or start and stop a game repeatedly. Compounding this is LucasArts' new Euphoria technology, using artificial intelligence, in which what were previously pre-programmed game events are now unscripted and random. "You'll never be able to predict exactly what will happen, no matter how many times you've experienced a certain scenario", they promise.<sup>22</sup> The composer, of course, cannot compose an infinite number of cues for a game, although games do now have an increasing number of cues to reduce the amount of looping.<sup>23</sup> Well aware that games have a reputation for repetitive, looping music—but trapped by the many hours of music required in a game—composers now commonly re-use cues in other areas of a game, to reduce the amount of unique cues needed, but without creating

a repetitive sounding score. This requires careful compositional planning, and often a reduction in dramatic melody lines so that the cue is less memorable.

Related to this temporal predicament is the concept of “listener fatigue”: games are designed to be played multiple times, and repeated listening can be tiring, especially if a player spends a long time on one particular area of the game. Games have begun to incorporate timings into the cues, so that if the player does get stuck on a level, the music will not loop endlessly, but will instead fade out. McConnell (1999) uses this technique extensively in *Grim Fandango*. He explains, “Another compositional approach to writing state (non-diegetic) music is to start with a short flourish, or ‘stinger’, and then end or fade the music and let the ambient sound effects take over”. Composer Marty O’Donnell elaborates in his discussion of the *Halo: Combat Evolved* score (Bungie Software 2001), “there is this ‘bored now’ switch, which is, ‘If you haven’t reached the part where you’re supposed to go into the alternative piece, and five minutes have gone by, just have a nice fadeout.’” (Cited in Battino & Richards 2005, 195).

### **Implications of Dynamic Sound on the Functions of Games Audio**

The dynamic, non-linear aspects of gameplay also impact the role and functions that audio serves in games. While games audio typically maintains all of the functions found in film or television sound, (see Berg 1973; Cohen 1999; Chion 1994; Gorbman 1987; Kozloff 1988; Lissa 1965; Manvell & Huntley 1975; Smith 1998), there are also some distinct differences in the ways audio functions in games. Depending on genre, platform and on the player’s familiarity with a game, some games can function without sound altogether, or with altered or substituted sound or music selected by the player. Games such as *Twisted Metal 4*, (989 Studios 1999) for instance, allow the player to remove the game’s music and replace it with their own chosen CD. *Grand Theft Auto: San Andreas* (Rockstar Games 2004) has “radio stations” available so that a player can select the kinds of music that they want to hear. Games for portable players such as Play Station Portable or Game Boy Advance are designed with the knowledge that these games are often played in the presence of other people and may require silence. As such, these devices often have much more limited capabilities for sound, and the audio is typically designed to serve a lesser role. Subsequently, the functions I describe below cannot be said to hold true for all games.

External to the games themselves is the economic impact that gaming has on various industries (and vice versa), including those of film and popular music. Increasingly, games are used as marketing tools and have become part of media franchises that may include film or television spin-offs. Games publisher Electronic Arts formed EA Trax as a marketing partner with labels “so as to not only find new music in all of our games, but hopefully create a music destination that gamers can rely on”.<sup>24</sup> Electric Artists, a music marketing agency, recently published a white paper on video games and music after surveying “hard-core gamers”, releasing such impressive statistics as : “40% of hard-core gamers bought the CD after hearing a song they liked in a video game”; “73% of gamers said soundtracks within games help sell more CDs”; and “40% of respondents said a game introduced them to a new band or song, then 27% of them went out and bought what they heard”.<sup>25</sup> With such impressive statistics, it appears indicative of a new relationship to popular music in general. After all, it is unlikely that equal numbers of people purchase a CD after hearing a song on the radio.

The interplay between audience and audio has other impacts on the ways in which popular music is consumed. In the case of *Guitar Hero* or *SingStar* (Sony 2004), there is a direct participatory and performance aspect to listening to the songs. In many music games, the player is placed in the role of the star, the performer, even if these games are meant primarily for home play.<sup>26</sup> *SingStar* advertises other mini-games on its web site, declaring “Live the life of a rock star, kick back in your limo, create your own look and more”.<sup>27</sup> A question arises, then, as to whether this fantasy changes the relationship between the audience and the music, or whether it has the same role as, for instance, playing “air guitar” and imagining being part of the band.

There are also aesthetic implications of the commercial aspects of games sound. A major repercussion of choosing licensed music is that there is limited adaptability or interactivity inherent in the music. Licensed songs are (primarily) designed as linear music, and therefore the placement of this music in games, (other than music games like *SingStar*), is generally limited to more linear aspects of gameplay (cut-scenes, title themes, credits, etc.), as is the *type* of game where such music may be appropriate. Certain types of games have become associated with specific genres of music, depending on who the target audience of the game is. Driving games, for instance, require “driving music” and are more likely to include popular songs with a repetitive “groove”.

Sergio Pimentel, who acquired the music for *Driver: Parallel Lines* (Reflective 2006), for instance, commented that he drove his car around listening to many types of music until he found the right “feel” for the game (2006), a combination of hip-hop and rock. To a certain degree, intertextual referencing and new semiotic connotations associated with the placement of linear music in games may occur.

The direct participation between a player and the audio takes on a new role in kinetic gestural games. These games are designed to have players directly physically participate and respond to the sound. Of course, such games are enjoyable—as is evidenced by their popularity—but the music is also sometimes intended as part of the edutainment role of some of these games (as in training basic motor skills in toddlers, for instance), or designed for aiding in physical fitness, such as *EyeToy Kinetic* (Sony 2005), which is clearly implicated in the marketing of these games. The sound in the case of kinetic games serves as a main motivating factor, arousing the player physically, and is also the part of the game on which the player must focus attention and with which the player must respond.

This changing role from passive to active listening is an important element of sound in games. A crucial semiotic role of sound in games is the preparatory functions that it serves, for instance, to alert the player to an upcoming event. Anticipating action is a critical part of being successful in many games, particularly adventure and action games. Notably, acousmatic sound—that is, sound with no clear origin visually—may inspire us to look to the direction of a sound, to “incite the look to go there and find out” (Chion 1994, 71, 85). This function, while present in films (even if we cannot force the camera to go in that direction, we mentally look there, argues Chion), is far more pronounced in games, as sound gives the player cues to head in a particular direction or to run the other way.<sup>28</sup> For instance, in *Zelda*, a giant rolling boulder is heard and grows in volume as it approaches—giving the player fair warning of an impending danger. The player is aware of the sound and will listen out for the boulder as they traverse that particular area in order to decide the appropriate time to move. As Marks explains, “Without [the audio], the player doesn’t have any foreshadowing and won’t know to take out their weapon or to get out of the room until it is too late. While this can lead to a learning experience for a player, repeatedly dying and having to start a level over can be frustrating enough to stop playing the game.” (2003, 190.)

Particularly important to games is the use of sound symbols to help identify goals and focus the player’s perception on certain objects. As Cohen describes, sound focuses our attention, as when a “soundtrack featuring a lullaby might direct attention to a cradle rather than to a fishbowl when both objects are simultaneously depicted in a scene” (2001, 258). In *Zelda*, the lesser enemies all have the same music, and beneficial items like gems or pieces of heart likewise all have the same or similar sounding cues. In other words, symbols and *leitmotifs* are often used to assist the player in identifying other characters, moods, environments and objects, to help the game become more comprehensible and to decrease the learning curve for new players. The use of recurrent musical themes can help to situate the player in the game matrix, in the sense that various locales or levels are usually given different themes. By listening to the music, the player is able to identify their whereabouts in the narrative and in the game. In the *Legend of Zelda*, for instance, musical themes play a key role, such as ‘Saria’s Song’, the theme taught to the main character by his friend, Saria. The recurrence of the theme in several places helps otherwise seemingly disparate scenes hold together and provides a degree of continuity across a game that takes weeks to finish, while reminding the player of previous scenes. It also serves to reinforce the theme in the player’s mind, so that when they learn to play the theme on the ocarina, it sounds familiar, and when they must recall the theme at specific points in the game, it will be likely remembered (Whalen 2004).

Music, then, can be used to enhance the overall structure of the game. These can include direct structural cues, such as links or bridges between two scenes, or which indicate the opening or ending of a particular part of gameplay. A drop to silence (the “boredom switch”) can also tell the player that they should have completed that segment of the game, and that the game is waiting for a player to overcome a particular challenge or exit the area. A pause or break in music can indicate a change in narrative, or, continuous music across disparate scenes can help to signal the continuation of a particular theme (Cohen 1999, 41). For games like *Vib Ribbon* (SCEI 1999), the music can literally *create* the structure of the gameplay. Released in Japan for the Playstation, the game allows the user to input his or her own music CDs, which then influences the game’s generation of level mapping. The game scans the user’s CD and makes two obstacle courses for each song (one easy and one difficult), so that the game is as varied as the music the player chooses. Although this case is fairly unique, the potential certainly exists for using music to influence structures or to personalise games, as has been explored in audio games, specifically designed for the visually impaired.

Equally important in reinforcing elements of gameplay is the dialogue, which can, for instance, disclose clues, or assign goals (Kozloff 2000, 5). For example, there are often hints and goals given in the dialogue in *Grim Fandango*. When Eva tells us she needs our teeth, for instance, we have to go and find an object that will suffice before we can progress in the game. Listening to dialogue, then, becomes a key element in solving the game. Sound and dialogue can likewise reveal details about places or characters—whether they are a friend or a foe, for instance, either by their musical accompaniment or by the accent/language/timbre of their voice and voice-over narrations can let us access a character’s thoughts and feelings (Kozloff 2000, 5). In *Grim Fandango*, the stereotyped accents play a key role in quickly understanding characters. South American Salvador is the revolutionary; the men criticizing capitalism in the jazz bar speak with a beat poet lingo, while the corporate boss, Don Copal, has a generic American accent. Changes in voice or accent are used to indicate other changes in gameplay: If the player chooses to have Manny take a drink of gold leaf liquor, his words slur for a while, providing a little added humour. While much of the verbal interplay between player and game has traditionally been text based, with lines selected or typed in by the player, future games will undoubtedly be more vocal, with players literally speaking to a game’s characters.

Part of the role of dialogue—and audio in general—in a game is the suspension of disbelief, adding realism or creating illusion. The illusion of being immersed in a three-dimensional atmosphere is greatly enhanced by the audio, particularly for newer games which may be developed in 7.1 surround sound. Even more simple stereo effects still have a considerable impact. In *Grim Fandango*, for example, the sound changes in the Calavera Café based on the character’s location, and follow the character’s proximity to the piano using stereo location and occlusion effects. In addition to spatial acoustics helping to create an environment, the music, dialogue and sound effects help to represent and reinforce a sense of location in terms of cultural, physical, social, or historical environments. For instance, the song playing at the Day of the Dead Festival in *Grim Fandango*, ‘Compañeros’, uses a blending of rural folk and Mexican mariachi band music, with trumpet, violin, guitar, *vihuela* and *guitarrón*. This is contrasted by ‘Swanky Maximino’, which plays in the office of Maximino, the gambling kingpin. Explains composer Peter McConnell,

I wanted to evoke the sounds of speakeasies of the Prohibition Era, those of the smaller “big” bands that played in clubs in such places as Harlem in the heyday of the Thompson gun-toting gangster... The music he would listen to, and the kinds of bands he would hire to play at the club, would reflect that sense of opulence as well. The main part of the music is inspired by recordings of the early Ellington band, when the Duke led the musical part of the Harlem Renaissance in the late twenties.<sup>29</sup>

While this function of game audio does not differ significantly from that of film, it must be recalled that a game may take thirty to forty hours to complete even when the “correct sequence” of events are known, and audio plays a crucial role in helping the player to recall places and characters, and to situate themselves in such a massive setting.

Another important immersive element Gorbman (1987) and Berg (1973) both discuss in relation to film is the historical function of covering distracting noises of the projector in the era of silent movies. A similar function may be attributed to games sounds created for an arcade environment. Arcade games have tended to have less polyphony and more sound effects and percussion, as part of the necessity of the environment, which meant that the games must be heard over the din to attract players. In consoles designed for home gameplay, music may mask the distractions of the computer fan, or sounds made by the surrounding environment (Cohen 1999, 41). I myself, for instance, crowded into a small apartment with a roommate, often turn the sound up on my headphones when working or gaming, to drown out telephone conversations, the snow plow outside, or other noises that serve to distract my attention. Although this has less of an effect on the audio than that of the arcade games, merely having a constant soundscape in a game can help the player to focus on the task at hand in a distracting environment.

Finally, adding to the immersive effects of gameplay is the communication of emotional meaning, which occurs in games audio in much the same way as in linear media. Here, a distinction must be made between communication of meaning through music, and mood induction: “Mood induction changes how one is feeling, while communication of meaning simply conveys information. One may receive information depicting sadness without him or herself feeling sad.” (Rosar in Cohen 2001, 42.) Mood induction and physiological responses are typically experienced most obviously when the player’s character is at significant risk of peril, as in the chaotic and fast boss music (the final major enemy of a level or series of levels in a game) in *Zelda*, for instance. In this

way, sound works to control or manipulate the player's emotions, guiding responses to the game. As many studies of video games violence show, many players are considerably emotionally involved in games.<sup>30</sup> Where games differ from linear media in terms of this relationship is that the player is actively involved in their character—there are consequences for actions that the player takes. If the character dies, it is the player's "fault", as this is not necessarily a pre-scripted event out of the player's control. I would argue that this creates a different, (and in some cases perhaps more intense) relationship between the player and the character(s).<sup>31</sup>

### **The Impacts of the Dynamic Aspects of Game Audio**

The ways in which audio is used and functions in a game, as we have seen, are impacted by many factors. I have shown elsewhere how the technological limitations of historical games audio consoles are critical to the understanding of the construction of sound in games as well as its semiotic implications (see Collins 2006b). Although recent games systems have overcome many of the constraints of the past, the issues of participation and non-linearity bring forth many unique complications which have both practical and theoretical implications.

There are clearly many interesting problems which face games composers and sound designers, and the issue of non-linearity in sound is crucial to our understanding of games audio, as the factors which influence composition and sound design clearly influence the sound's aesthetics. In considering games audio, there are new consumption modes, new production modes and new listening modes which distinguish themselves from linear media. I have discussed the ways in which transitions impact a game, and how an inappropriate transition may result in a loss of the immersive characteristics of a game. Composers have begun to use cue fragments to adapt to this problem, but the issue calls forth aesthetic, as well as technological complications. To return to our metaphor of the metro station, not only may the passenger jump at any time to another car, but the emotional impact of that jump must also be considered in the transition. For example, if the first car was full of happy people laughing, even a cross-fade to an angry car may leave the passenger emotionally confused. The music, then, must be scored in subtle ways to build up to any dramatic emotion. Without a click-track to follow, the best the composer can hope for is to score general moods rather than direct timings to actions. Timbres are more limited, as abrupt changes from, say, trumpet jazz to heavy metal guitar are going to be awkward.

As discussed, popular music is playing an increasingly more significant role in games. Statistics suggested that gamers purchase a considerable proportion of music that they hear in a game. As such, this raises the possibility that repetition—or the associated enjoyment with play that occurs in a game—may influence the decision to purchase. More noteworthy, perhaps, is the fact that in many games it is unlikely that the player will hear the entire song, but instead may hear the first opening segment repeatedly, particularly as they try to learn a new level. The possible implication of this may be that, in the future, songs selected for games may be created with the idea that the player may not ever hear more than the first part of the song, thus changing the ways in which popular music is constructed.

Unlike in most film and television—in which music is usually designed to be absorbed on a purely subconscious level as "unheard melodies" (Gorbman 1987),<sup>32</sup> audio in games typically serves more of an active function, and at times must be carefully listened to in order to make the correct response. Depending on the type of game, turning off the sound, or failure to pay attention to the sonic cues, can lead to loss of points or even a character's demise. Ultimately, sound functions and is consumed in games in different ways than those of other media, and as such, our relationship to sound in games must be theorised differently from ways of the past.

I have shown that the issue of participation may require new semiotic models to understand how meaning is communicated in games, introducing the idea of "participatory supplementary connotations", since existing models rely on the notion that the receiver has no input on the transmission. There are further questions relating to reception that are raised by games, particularly the non-linear aspects which may indicate changing conceptions of time in our culture. Kramer, for instance, suggests that "just as discontinuous life-styles are becoming the norm and just as the continuity in modern time-arts is of a very different order than that in classical time-arts, so non-linear modes of experiencing pre-contemporary music are contrasting with, and even supplanting, our traditional and in a sense nostalgic well-ordered time-experiences" (1973, 124.) Kramer links the growth in non-linearity in Western music to both technological advances and social discontinuity. Perhaps the non-linear aspects of games audio are indicative of a wider cultural paradigm shift, as previous changes in music have also reflected their times (see Atalli 1985, Tagg 1996, etc.). Does the reception of non-linear games audio

correspond with other contemporary non-linear audio, or does the experience of gaming change or influence the reception of non-linearity?

There are clearly issues of temporality in games music that require further exploration but are outside the scope of this paper. The length of gameplay, as discussed, is different from that of film—it could be one minute, or twelve hours, for instance, that a player chooses to engage in that mode. It may be that the same segments of music in the same order are only heard once in a game or it may be that one piece is looped for hours. The resultant implications for the composer are a reduction of dramatic melodies, but what impacts do cue fragments—and the mutability of the music—have on a player? With the adaptability of cue fragments, the player may hear what is essentially one long song for the length of their gameplay—how does the lack of a structured beginning, middle and end affect the listener’s reception of the music? If all of our paradigms for thought are based on linearity, what are the wider implications of having the ability to conceive of ideas in a non-linear framework?

It may be necessary to draw on theories of drones and the loss of time sometimes associated with their reception,<sup>33</sup> or on some of the theories of new temporality and non-linearity in music discussed by such authors as McClary (2000), Tagg (1996), Kramer (1973 and 1981), etc. Kramer’s notion of “vertical time” may be useful in understanding games audio, as he suggests vertical time “consists of relationships between ever-present layers of the dense sound world, whereas form in linear music consists of relationships between successive events” (1981, 551). Kramer argues,

The result is a single present stretched out into an enormous duration, a potentially infinite “now” that nonetheless feels like an instant...it does not begin but merely starts, does not build to a climax, does not purposefully set up internal expectations, does not seek to fulfill any expectations that might arise accidentally, does not build or release tension, and does not end but simply ceases. (1981, 549.)

Likewise, repetition, looping and its implications have been explored further by many authors (Garcia 2005; Katz 2004; Middleton 1996; Spicer 2004; Stillar 2005, etc.), although whether the loop in context with the visual image and dynamic nature of the gameplay is received differently than loops in linear popular music still remains an unanswered question. Garcia (2005) explains how the repetition of seemingly short musical units can generate pleasure over extended periods. Has games audio helped looping to become more acceptable in popular music, by changing the way it is received and listened to? These ideas and concepts are not limited to sound. In a game like *Zelda*, the player has the option to change perspectives and point of view in the game. While the default is a third-person point of view, the player can opt for first-person point of view, as well as change the angle of viewing, in a sense becoming the director and influencing the way visual imagery is produced and consumed in games. Just as MTV helped to influence new styles in film editing,<sup>34</sup> we need to now consider video games as a major cultural influence with implications on both practice and theory in popular music and cultural theory in general.

## List of References

- Attali, Jacques 1985: *Noise: The Political Economy of Music*. University of Minnesota Press, Minnesota.
- Bajakian, Clint; David Battino & Keith Charley 2003: Group Report: What is Interactive Audio? And What Should It Be? The Eighth Annual Interactive Music Conference Project Bar-B-Q. Internet Source, available <http://www.projectbarbq.com/bbq03/bbq03r5.htm> (12.3.2006).
- Battino, David & Kelli Richards 2005: *The Art of Digital Music: 56 Visionary Artists and Insiders reveal their Creative Secrets*. Backbeat Books, San Francisco.
- Berg, Charles Merrell 1973: *An Investigation of the Motives for and Realization of Music to Accompany the American Silent Film, 1896-1927*. PhD Thesis, Department of Speech and Dramatic Art, University of Iowa.
- Bessell, David 2002: What's That Funny Noise? An Examination of the Role of Music in *Cool Boarders 2*, *Alien Trilogy* and *Medieval 2*. In King, George & Krzywinska, Tanya (eds.) *ScreenPlay: Cinema/videogames/interfaces*. Wallflower, London.
- Bordwell, David & Thompson, Kristin 1990: *Film Art: An Introduction*. McGraw-Hill, New York.
- Calvert, S.L., & Tan, S. 1994: Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation. *Journal of Applied Developmental Psychology*, 15, 125–139.
- Branigan, Edward 1984: *Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film*. Moulton, Berlin.
- Chion, Michel 1994: *Audio-Vision: Sound on Screen*. Columbia University Press, New York.
- Cohen, Annabel J. 1998: The Functions of Music in Multimedia: A Cognitive Approach. In Yi, S.W. (ed.): *Music, Mind & Science*. Seoul University Press, Seoul, 40-68.
- Cohen, Annabel J. 2001: Music as a Source of Emotion in Film. In Juslin, Patrick N. & Sloboda, John A. (eds.) *Music and Emotion: Theory and Research*. Oxford University Press, Oxford, 249-279.
- Collins, Karen 2002. *The Future is Happening Already: Industrial Music, Dystopia, and the Aesthetic of the Machine*. PhD thesis, Institute of Popular Music, University of Liverpool, Liverpool.
- Collins, Karen 2006a: Loops and Bloops: Music on the Commodore 64. *Soundscapes: Journal of Media Culture*. Volume 8: February. Internet source, available <http://www.icce.rug.nl/~soundscapes> (3.3.2006).
- Collins, Karen 2006b: Flat Twos and the Musical Aesthetic of the Atari VCS. *Popular Musicology Online*, Issue 1. Internet source, available <http://www.popular-musicology-online.com> (11.06.2006).
- Cook, Nicholas 2000: *Analysing Musical Multimedia*. Oxford University Press, Oxford.
- Curtis, Scott 1992: The Sound of Early Warner Bros. Cartoons. In Altman, Rick (ed.): *Sound Theory Sound Practice*. Routledge, London, 191-203.
- Dancyger, Ken 2001: *The Technique of Film and Video Editing: History, Theory, and Practice*. Focal Press, London.
- Fay, Todd M.; Selfon, Scott & Fay, Todor J. 2004: *Directx 9 Audio Exposed: Interactive Audio Development*. Wordware Publishing, Texas.
- Folmann, Troels 2006: *Tomb Raider Legend: Scoring a Next-Generation Soundtrack*. Conference presentation for the Game Developer's Conference 2006. San Jose, California.
- Forlenza, Jeff & Stone, Terri 1993: *Sound For Picture: An Inside Look at Audio Production for Film and Television*. Hal Leonard Publishing Corporation, Emeryville, CA.
- Garcia, Luis-Manuel 2005: On and On: Repetition as Process and Pleasure in Electronic Dance Music. *Music Theory Online* Volume 11, Number 4. Internet source, available <http://mto.societymusictheory.org/issues/mto.05.11.4/mto.05.11.4.garcia.html> (12.2.2006).
- Gentile, D.A.; Lynch, P.L.; Linder, J.R.; & Walsh, D.A. 2004: The Effects of Violent Video Game Habits on Adolescent Hostility, Aggressive Behaviors, and School Performance. *Journal of Adolescence* Issue 27, 5–22.
- Gorbman, Claudia 1987: *Unheard Melodies: Narrative Film Music*. Indiana University Press, Bloomington.
- Griffin, Donald S. 1998: Musical Techniques for Interactivity. *Gamasutra*. May 1, 1998, Vol. 2: Issue 18. Internet source, available [http://gamasutra.com/features/sound\\_and\\_music/19980501/interactivity\\_techniques\\_03.htm](http://gamasutra.com/features/sound_and_music/19980501/interactivity_techniques_03.htm) (12.3.2006).
- Grim Fandango Files 1998: Internet source, available [http://www.lucasarts.com/products/grim/grim\\_files.htm](http://www.lucasarts.com/products/grim/grim_files.htm) (12.4.2006).
- Hainge, Greg 2004: The Sound of Time is not *tick tock*: The Loop as a Direct Image of Time in Noto's *Endless Loop Edition* and the Drone Music of Phill Niblock. *Invisible Culture: Electronic Journal for Visual Culture*, Issue 8. Internet source, available [http://www.rochester.edu/in\\_visible\\_culture/Issue\\_8/hainge.html](http://www.rochester.edu/in_visible_culture/Issue_8/hainge.html) (12.4.2006).
- IGN's Top 100 Games. Internet source, available <http://top100.ign.com/2005/index.html> (5.3.2006).
- Irwin, A.R. & Gross, A.M. 1995: Cognitive Tempo, Violent Video Games, and Aggressive Behavior in Young Boys. *Journal of Family Violence* Issue 10, 337–350.
- Kassabian, Anahid 2003: The Sound of a New Film Form. In Inglis, Ian (ed.) *Popular Music and Film*. Wallflower, London, 91-101.

- Katz, Mark. 2004: *Capturing Sound: How Technology has Changed Music*. University of California Press, Berkeley.
- Kozloff, Sarah 1988: *Invisible Storytellers: Voice-over Narration in American Fiction Film*. University of California Press, Berkeley.
- Kozloff, Sarah 2000: *Overhearing Film Dialogue*. University of California Press, Berkeley.
- Kramer, Jonathan D. 1973: Multiple and Non-Linear Time in Beethoven's Opus 135. *Perspectives of New Music*, Vol. 11, No. 2, 122-145.
- Kramer, Jonathan D. 1981: New Temporalities in Music. *Critical Inquiry*, Vol. 7, No. 3, 539-556.
- Land, Michael Z. & Peter N. McConnell 1994: Method and Apparatus for Dynamically Composing Music and Sound Effects using a Computer Entertainment System. United States Patent Number 5,315,057.
- Lissa, Zofia 1965: *Ästhetik der Filmmusik*. Henscherverlag, Berlin. Translated and summarised by Tagg, Philip. Internet source, available <http://www.mediamusicstudies.net/tagg/udem/musingmot/filmfunx.html> (2.4.2006).
- Manvell, Roger & Huntley, John 1975: *The Technique of Film Music*. Focal Press, New York.
- Marks, Aaron 2001: *The Complete Guide to Game Audio: For Composers, Musicians, Sound Designers, and Game Developers*. CMP Books, California.
- McClary, Susan 2000: Temporality and Ideology: Qualities of Motion in Seventeenth-Century French Music. *Echo: A Music-Centered Journal*. Volume 2 Issue 2, Fall 2000. Internet source, available <http://www.humnet.ucla.edu/echo> (3.3.2006).
- McConnell, Peter 1999: The Adventures of a Composer Creating the Game Music for Grim Pandango (*sic*). *Electronic Musician*. Internet Source, available [http://emusician.com/mag/emusic\\_adventures\\_composer\\_creating/index.html](http://emusician.com/mag/emusic_adventures_composer_creating/index.html) (20.5.2006).
- Middleton, Richard 1996: Over and Over: Notes towards a Politics of Repetition. Surveying the Ground, Charting Some Routes. Conference paper presented at *Grounding Music for the Global Age*, Berlin. Internet source, available <http://www2.hu-berlin.de/fpm/texte/middle.htm> (2.4.2006).
- Moore, Shaun 1993: *Interpreting Audiences: The Ethnography of Media Consumption*. Sage, London.
- Pimentel, Sergio 2006: Music Acquisition for Games. Conference presentation for the Game Developer's Conference 2006. San Jose, California.
- Seflon, Scott 2006: Audio Boot Camp: Music for Games. Conference presentation for the Game Developer's Conference 2006. San Jose, California.
- Smith, Jeff 1998: *The Sounds of Commerce: Marketing Popular Film Music*. Columbia University Press, New York.
- Spicer, Mark 2004: (Ac)cumulative Form in Pop-Rock Music. *Twentieth-Century Music* 1/1, 29-64.
- Stillar, Glenn 2005: Loops as Genre Resources. *Folia Linguistica* XXXIX/1-2, 197-212.
- Stockburger, Axel 2005: The Game Environment from an Auditive Perspective. Internet source, available <http://www.audiogames.net/ics/upload/gameenvironment.htm> (4.6.2006).
- Tagg, Philip 1996: Understanding 'Time Sense'. Internet source, available <http://www.mediamusicstudies.net/tagg/articles/timesens.html> (20.4.2006).
- Tagg, Philip 1999: Introductory Notes to the Semiotics of Music. Postgraduate class handout, Institute of Popular Music, University of Liverpool, Liverpool.
- Tagg, Philip 2002: Notes on how Classical Music Became 'Classical'. Internet source, available <http://tagg.org/teaching/classclassical.pdf> (20.4.2006).
- Video Games of Note. Internet source, available [http://www.electricartists.com/videogamesofnote\\_whitepaper\\_0314.pdf](http://www.electricartists.com/videogamesofnote_whitepaper_0314.pdf) (22.1.2005).
- Whalen, Zach 2004: Play Along: An Approach to Videogame Music. *Game Studies: The International Journal of Computer Game Research*, 4/1. Internet source, available <http://www.gamestudies.org/0401/whalen> (3.3.2006).
- Whitmore, Guy 2003: Design with Music in Mind: A Guide to Adaptive Audio for Game Designers. *Gamasutra*. Internet source, available [http://www.gamasutra.com/resource\\_guide/20030528/whitmore\\_01.shtml](http://www.gamasutra.com/resource_guide/20030528/whitmore_01.shtml) (2.4.2006).
- Wolf, Mark J.P. (ed.) 2001: *The Medium of the Video Game*. University of Texas Press, Austin.
- Wolf, Mark J.P. & N. Bernard Perron (eds.) 2003: *The Video Game Theory Reader*. Routledge, New York.

#### **Audio-Visual References**

- Asheron's Call 2: The Fallen Kings*. Turbine Entertainment 2003
- Donkey Konga*. Namco 2003
- Driver: Parallel Lines*. Sony 2006
- Eye Toy: Kinetic*. Sony 2005
- Grand Theft Auto: San Andreas*. Rockstar 2004
- Grim Fandango*. LucasArts 1998

*Guitar Hero*. Red Octane 2005  
*Halo: Combat Evolved*. Bungie Software 2001  
*Legend of Zelda: The Ocarina of Time*. Nintendo 1998  
*Legend of Zelda: The Twilight Princess*. Nintendo 2006  
*Legend of Zelda: The Wind Waker*. Nintendo 2003  
*Russian Squares*. Microsoft 2002  
*Sims*. Maxis 2000  
*SingStar*. Sony 2004  
*Super Mario Brothers*. Nintendo 1985  
*Tomb Raider: Legend*. Eidos 2006  
*Twisted Metal 4*. 989 Studios 1999  
*Vib Ribbon*. SCEI 1999

---

<sup>1</sup> I would like to acknowledge the kind guidance and advice of the editors and other readers of preliminary drafts of this paper, particularly Paul Théberge and John Richardson. By “video” games I refer to all digital games, including those on computers, consoles, mobile phones, etc.

<sup>2</sup> Tagg 2002, 1.

<sup>3</sup> This is not to suggest that all music—or even all Western music—is linear. See for instance, Kramer 1981 for further discussion of non-linearity in music. I have borrowed the train metaphor from Griffin 1998.

<sup>4</sup> Comparisons with film music have, unfortunately, fallen prey to implying games (at least, those of the past) are inferior to that of film. For example, David Bessell tells us that a particular game gives us “a rather one-dimensional impression in comparison to most film scores” 2002, 138. Bessell does acknowledge that there are difficulties in scoring dynamically, but the solutions he suggests—such as the approach taken by Boulez in using random sequences (2002, 142)—were ideas that had, in fact, been used in games going back to at least those of the Commodore 64 in the early 1980s (see Collins 2006a). It is quite common for games to introduce randomization, though perhaps in subtle ways. For instance, Marty O'Donnell and Jay Weinland of Bungie Studios inform us, “If you sit in one place in *Halo* listening to a piece of music, you will notice that it never plays back exactly the same way twice due to the randomization of the permutations. The music was edited so that the ‘in’ soundfile plays seamlessly into any of the main loop permutations, and the main loop permutations can play seamlessly into the ‘out’ soundfile when the piece is stopped by the script. These main loop permutations contain unique elements, so as they are randomly played back, the piece takes on a new flavor each time you hear it” (in Fay et al. 2004, 421).

<sup>5</sup> See for instance Wolf 2001 or Wolf & Perron 2003. A great majority of work into games has centered on their violence, but other approaches in the field of ludology have been increasing in the last decade.

<sup>6</sup> Such as Bessell 2002, Stockburger 2005, and Whalen 2004. One issue facing game audio musicologists has been transcription and copyright. With many old games companies now defunct—and many old games simply not giving credit to composers—it has been difficult to publish in some areas.

<sup>7</sup> Emulators, which can re-create the N64 platform on a home PC, are downloadable from the Internet, as are ROMs, the games themselves. Nintendo makes it clear on their website that these devices are illegal, and did not respond to my requests for special permission to use them for research purposes. Nevertheless, N64 consoles and games are available quite affordably on E-Bay and in used games stores.

<sup>8</sup> Sound design and production by Michael Land, Clint Bajakian, Nick Peck, Andy Martin, Jeff Kliment, Michael McMahon, Julian Kwasneski, and Kristen Becht. It is worth noting that such a large audio group is rare, and is typically only seen at larger studios. In many smaller games companies, the composers also serve as sound (effects) designers. With larger companies, such as Nintendo or LucasArts, these roles are often divided and sometimes filled by several people.

<sup>9</sup> According to IGN's Top 100 Games. IGN Entertainment is part of Fox Interactive Media, and is one of the largest reviewers of interactive media. Internet source, available <http://top100.ign.com/2005/index.html> (5.3.2006).

<sup>10</sup> See for instance [http://www.songbirdocarina.com/zelda\\_songs.html](http://www.songbirdocarina.com/zelda_songs.html). See also Lark in the Morning, <http://www.larkinthemorning.com> (6.4.2006).

<sup>11</sup> It may be useful to refer also to other approaches to sound theory. Scott Curtis, for instance, suggests a division of “isomorphic” and “iconic” uses of sound to replace the diegetic/non-diegetic dichotomy. Isomorphic “refers to the close matching of the image and sound—that is, a relationship based on rhythm in both the action and the music” (1992, 201), whereas iconic sounds have an analogous relationship between sound and image—“visual elements and the timbre, volume, pitch, and tone of the accompanying sound” (1992, 202). However, this distinction does not help us to understand the ways that the audience interprets the sounds in relation to the fictional on-screen world, and so is unhelpful to the present argument.

<sup>12</sup> And could be sub-divided further, for instance, as Michel Chion has elaborated in his discussion of the diegetic/non-diegetic divide (1994), or Gorbman's “metadiegetic” subjective sound categories (1987, 450), Branigan's “intradiegetic” (1984, 68) and Bordwell & Thompson's “internal diegetic” (1990, 256), etc. Although the use of the division of diegetic/non-diegetic is called into question in contemporary film studies because of its inability to deal with these other categories of

---

sound, I do not wish to further complicate the issue here. My point is that the relationship of the audio to the player and to their character in games is different than in film because of the participatory nature of games.

<sup>13</sup> Cut-scenes require “a dramatic score to grab a player’s attention and motivate them to make it through the next level to see the next movie. These offerings often serve as a reward to the player, where they are able to watch the on-screen action and inwardly pride themselves for making it happen.” See Marks 2001, 190.

<sup>14</sup> See for instance, Moores 1993.

<sup>15</sup> See also Tagg 1999, 11. Although the concept of supplementary connotation may be latent in Tagg’s concept of codal interference, it is not clearly conceptualized: “when transmitter and receiver share the same basic store of musical symbols but totally different sociocultural norms and are basically ‘understood’ but that ‘adequate response’ is obstructed by other factors, such as receivers’ general like or dislike of the music and what they think it represents, or by the music being recontextualised verbally or socially”.

<sup>16</sup> Although the term “score” is more commonly used to refer to printed musical compositions, it is the term commonly used by those in the games industry to refer to music composed for games. Increasingly, interactive audio compositions are being scored, orchestrated by professionals, and recorded by choirs or orchestras.

<sup>17</sup> I rely on personal experience and the experience of some of my students for this observation.

<sup>18</sup> Cue fragments are also referred to as “audio chunks” amongst games composers. Troels Folmann calls them “microscores”. Folmann’s score for *Tomb Raider: Legend* (Eidos 2006), for instance, contained over three hours of music in more than five hundred individual cues. Folmann 2006.

<sup>19</sup> The book provides detail and examples from the composition. I use the ‘Gravity Ride’ track for this discussion. The game is part of the XP Plus package offered by Microsoft.

<sup>20</sup> The layout of my transcription is borrowed from Garcia 2005. The blocks of bars are laid out in much the same fashion as sequencing software. The passage of time is followed from left to right, with each block representing one bar of music. Each vertical layer in the stack of blocks represents a new instrument sound.

<sup>21</sup> There are of course other reasons for this sound/music disintegration. Recent technology has made this far easier than in the past, for instance. I would speculate that one of the most significant contributing factor to this idea is the changing role of the *sound effects editor* to *sound designer* in the last few decades and the subsequent more creative uses of non-musical sound in film. Blair Jackson, sound designer for television show *Twin Peaks*, for instance, described his approach to sound in one scene, “they went into the scene with music, then favored the effects, and then went out of it musically. It happened so seamlessly you didn’t realize it wasn’t one continuous flow of sound. It was all part of the ‘musical score.’” See Forlenza & Stone 1993, 110.

<sup>22</sup> From the LucasArts press release, *Indy Game Features Euphoria Technology*. April 27, 2006. Internet source, available: <http://www.indianajones.com/videogames/news/news20060427.html> (05.05.2006).

<sup>23</sup> Algorithmic generation of cues is likely a major area to be reconsidered in the coming years.

<sup>24</sup> EA Trax Interview, internet source, [http://www.music4games.net/f\\_eatrax.html](http://www.music4games.net/f_eatrax.html) (22.1.2005).

<sup>25</sup> See “Video Games of Note” on [http://www.electricartists.com/videogamesofnote\\_whitepaper\\_0314.pdf](http://www.electricartists.com/videogamesofnote_whitepaper_0314.pdf) (22.1.2005).

<sup>26</sup> It is interesting to note that the original version of *Singstar* featured a single-player mode, but this function was dropped in later releases of the game, indicating a clear preference for group play.

<sup>27</sup> See the *SingStar* website, <http://www.singstargame.com/> (28.5.2006).

<sup>28</sup> Stockburger 2005 provides a fairly extensive analysis of what he terms the “spatialising functions in the game environment”, focusing on what he has called the “dynamic acousmatic function”, which, he elaborates, is distinguished from film by the “kinaesthetic control over the acousmatisation and visualisation of sound objects in the game environment”.

<sup>29</sup> Internet source, available [http://www.lucasarts.com/products/grim/grim\\_files.htm](http://www.lucasarts.com/products/grim/grim_files.htm) (5.6.2006).

<sup>30</sup> See Gentile et al, 2004, Irwin & Gross 1995, or Calvert & Tan 1994, for instance. Of course, emotional involvement in games is not limited to violence. The *Guardian* had an article in which it discussed the emotional reaction to games, and reported later, “In response to the article, many gamers confess to having cried when Aeris died in *Final Fantasy VII*”. See the games blog, [http://blogs.guardian.co.uk/games/archives/2005/09/19/im\\_crying\\_here.html](http://blogs.guardian.co.uk/games/archives/2005/09/19/im_crying_here.html) (05.05.2006). As a personal anecdote, I can recall a time when I first played *The Sims* (Maxis 2000). It surprised me that in a very short time I began to think of my life in *Simean* terms—adding friends, collecting objects, etc. I remarked on this somewhat disturbing realization to a friend, who commented that after playing *Roller Coaster Tycoon 2* (Infogrames 2002) she, too, had a similar experience looking at real-world buildings and trying to judge if they would fit into the game world she was constructing. Similarly, upon discussing a particularly stressful time recently with a games sound designer friend of mine, he remarked, “you just need to find some power-ups”. Concepts introduced in gaming had entered and become useful in our lives. While we could differentiate between the “real” world and the world of the game, we had begun to think of our lives as if we were living out a video game experience.

<sup>31</sup> However, the active involvement of the player with a game also has other repercussions that can lessen the immersive quality, such as the player’s fumbling with the controller, for instance. Even after playing a game for several hours, for example, I find that I forget what function some of the buttons have, and this can cause the immersive effects of the game to lessen.

<sup>32</sup> There are of course exceptions, such as in commercials, in which a track becomes part of a product’s branding. See Cook 2000, 21.

<sup>33</sup> See for instance Hainge 2004.

<sup>34</sup> See Dancyger 2001.