Chapter 7

Gameplay, Genre, and the Functions of Game Audio

As shown in the first chapters of this book, game audio has been significantly affected by the nature of the technology (in terms of hardware, software, production, and distribution technology) and by the nature of the industry (in terms of design, production, distribution, and marketing). Audio in games is also, of course, affected by the nature of games themselves, in terms of genre, narrative, the participatory aspects of games, and the functions that audio must fulfill. As shown in the previous chapter, one of the ways to reduce the risks connected with the highly competitive nature of the games industry is through an association with star talent and licensed intellectual property.

As Martin Campbell-Kelly (2004, p. 45) notes in his history of the software industry, before becoming associated with movie tie-ins and celebrities, the first conscious risk-management strategy for the games industry was genre publishing. He argues that the point of genre publishing was that certain categories of games appealed to specific user-groups and therefore "had less need to overcome market inertia," noting that "From their earliest days, videogames had been classified as racin', fightin' or shootin'" (ibid., p. 281). Genre in games is particularly important in that it helps to set the audience's expectations by providing a framework for understanding the rules of gameplay, thereby not only appealing to specific target markets, but also reducing the learning curve of the game. Within the context of genre, many games function similarly in terms of their user interfaces, their button functions, and their rules of gameplay. These similarities were particularly important in the arcades, when players would walk away from a game before investing their money in a game that was too complicated to learn. Genres are also constrictive in many ways, however, and consumers may lament formulaic game designs. It has even been argued that it was such a formulaic approach to games construction that contributed to the brief collapse of the games industry in the mid-1980s (see Kline, Dyer-Witherford, and de Peuter 2003, pp. 104–108).

Some recent genre-defying games have seen remarkable sales, suggesting that genre is not central to the success of a game. *Electroplankton* (Nintendo, 2005), for instance, for Nintendo DS, in many ways is not even a ‘‘game’’ as traditionally defined. Nintendo president Satoru Iwata has described it as ‘‘a means for people to see sound’’ (in Totilo 2006). In the game, there are ten gameplay modes containing various types of plankton. Through making sounds by blowing into the DS microphone or by using the DS touch-screen stylus, the player makes the plankton move about and make musical sounds. In other words, by manipulating the plankton, the user can essentially make music, integrating notions of games with music production technology. Questions arise, then, as to whether or not such a program constitutes a ‘‘game’’ in the traditional sense. Similar questions have arisen out of recent virtual worlds like *Second Life* (Linden Lab, 2003): ‘‘It has no objectives, and the content is entirely user-created. But if *Second Life* isn’t a game, what exactly is it?’’ asks one journalist (Kalning 2007). The success of such games suggests that users are open to new forms and new genres, although it is not yet apparent if their popularity relies strictly on their novelty, or if knock-offs will enjoy similar success.

As with the amorphous nature of popular music or film genres, there does not seem to be any real consensus on game genres; their boundaries shift with each new release. Stephen Poole (2000, p. 44) distinguishes between nine genres, while Mark J. P. Wolf (2002) describes some forty-two different genres (see also Kerr 2006, pp. 38–40). As my intent is to highlight the differences that genre may sometimes place on games in terms of the audio, I will focus on the most popular genres, as defined by the current industry categorizations outlined by the Entertainment Software Association (ESA).[1] The ESA now refers to what it calls ‘‘supergenres,’’ presumably since there are too many games that cross over between smaller subgenres or defy genre definition altogether. They list these supergenres as fighting, role-playing, children and family entertainment, action, adventure, strategy, role-playing, shooter, racing, and sports. Game genres such as these are distinguished through several main characteristics, such as narrative, representational rules (the ‘‘realism’’ of games varies by genre, for instance), rules of gameplay, and types of interactivity or interface with the player. Each of these characteristics has a distinct impact on the ways in which audio functions in terms of its relationship to the player, and to the game’s narrative or diegesis.

DEGREES OF PLAYER INTERACTIVITY IN DYNAMIC AUDIO

The notion of diegesis, borrowed from film and drama studies, is perhaps ill suited to games. Nevertheless, it provides a useful notion with which we can discuss the different degrees of interaction between player/viewer and the on-screen content, and highlight some of the distinctions between the linear qualities of film audio and the nonlinear qualities of games.[2]

Dynamic audio complicates the traditional diegetic–nondiegetic division of film sound. The unique relationship in games posed by the fact that the audience is engaging directly in the sound playback process on-screen (discussed further below) requires a new type of categorization of the sound–image relationship. Game sound can be categorized broadly as diegetic or nondiegetic,[3] but within these broad categories it can be separated further into nondynamic 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 cinematics, which are fixed in a linear fashion, the degree of dynamic activity in a game is sometimes fluid, posing further difficulty in the classification of the sounds. For instance, in adventure game *The Legend of Zelda: Ocarina of Time'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 when a player enters a building or encounters an enemy. Though 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 now become adaptive. In other words, a cue that is interactive or adaptive at one point in the game does not necessarily remain so throughout the entire game. Similarly, in online RPG *Asheron's Call 2: The Fallen Kings* (Turbine Software, 2003), the nondiegetic music that plays in the background of scenes as underscore 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 nondynamic to interactive, but it has also gone from nondiegetic to diegetic. As such, then, although I have distinguished the levels of sound here, they must be viewed as fluid, rather than fixed, for many types of audio cues.

The most basic level of nondiegetic audio for games is the nondynamic linear sounds and music found most frequently in the introductory movies or cinematics. In these cases, the player has no control over the possibility of interrupting the music (short of resetting or turning off the game).[4] In the introduction to *The Legend of Zelda: Ocarina of Time* (hereafter just *Ocarina*), for instance, a short dream sequence movie is played, explaining the plot. If the player does not start the game (leading to further cinematics), the entire introduction sequence loops. At key points in the game, a preset cinematic loads, leading us to the next stage in the plot. In the LucasArts adventure game *Grim Fandango* (1998), the player's character Manny meets with Salvador, the revolutionary, in his underground hideout to conspire to expose the inequities of the corporation for which they work. When Manny gives Salvador the molded impression of his teeth (necessary for access to the building), a cinematic ends that stage of the game (El Marrow) and leads us to the next location (the Petrified Forest). The music during this intermission cinematic 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, *nondynamic, nondiegetic music*.

Nondiegetic audio can also contain various levels of dynamic activity. *Adaptive nondiegetic* 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 *Ocarina* 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 retrigger these events (except by waiting for another day to pass); the sounds are triggered by a timer set in the game engine. *Interactive* nondiegetic 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 *Ocarina*, 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 nondynamic, adaptive, or interactive. In *nondynamic diegetic audio*, the sound event occurs in the character's space, but the character has no direct participation with it. These sounds of course occur in cut-scenes, but they 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 nondynamic. Diegetic sounds can also be adaptive and interactive. To return to the night–day division of time in *Ocarina*, 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 nighttime, 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, and the player's character can directly interact with them. 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 preselected cue. More commonly, interactive diegetic sounds are sound effects, for instance, the sound Link's sword makes when cutting, or the sound of characters' footsteps.

Finally, a level of even more direct audio interaction is that of *kinetic gestural interaction* in both diegetic and nondiegetic 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 those times when a player may physically, gesturally mimic the action of a character, dancer, musician, etc. in order to trigger the sound event, most commonly seen in rhythm-action games. 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 of the traditional joystick/controller that is included with the game's platform), although this has changed since the release of Nintendo's Wii controller in 2006, which has made 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.

THE FUNCTIONS OF GAME AUDIO

The varying degree with which the player can interact with the sound suggests that audio can serve a wide variety of functions. 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 for portable players are often designed with the knowledge that these games tend to be played in the presence of other people and may require silence, as shown in the Nokia quote on page 78, which reminded programmers that ''The game should be playable without the sounds'' (Nokia 2005). More significantly, Microsoft has insisted that music in every Xbox360 game should be replaceable with the user's own music files (Harlin 2007, p. 53). In some ways, the reaction of the game audio community has been that in order to combat the implied ''uselessness'' or substitutability of games music by these requirements is to make the audio a more integral part of the games experience to ensure that the user does not switch off the sound.

Recent statistics suggest that game audio plays a significant role in consumer preference in product selection, and that audio is viewed as an important component of games (Bush et al. 2007). I have also shown some examples of games in which audio plays a specific role in the narrative or interface, such as stealth games and rhythm-action games. Indeed, it is quite evident that in many cases turning off audio would lead the player into peril. In the *New Super Mario Bros.* for the Nintendo DS (Nintendo, 2006), for example, enemies jump and fly in time to music, so listening to the sound signals to the player when to make his or her moves. Such use of audio indicates that game sound can be a significant element of gameplay in at least some cases, and that it can function in many ways (see also Jørgenson 2008).

Although game audio typically maintains all of the functions found in film or television sound (see Berg 1973; Chion 1994; Cohen 1999; Gorbman 1987; Kozloff 1988; Lissa 1965; Manvell and Huntley 1975; Smith 1998), there are also some distinct differences in the ways in which audio functions in games. There are some functions of film sound, for instance, that are not present to any significant degree in games (with the exception of cinematics). Consider for instance a chase scene in a horror film. There may be many close-up shots, pans, slow-motion shots of feet running, and so on—in other words, the editing of the cue may be quite slow when watched without sound. Here, the music and sound add the energy and pacing to make these edits flow together in a fast-moving sequence. In other words, the sound and music in film is often closely tied to the *edit*. That is, either the film is cut (edited) to the music, or the music is edited to fit the cut. With games, however, most action takes place in real time. Although scripted slow-motion shots do exist, image and music cannot synchronize as closely in games as they can in film, because of the unpredictable temporal aspects of games. Nevertheless, most other functions of film sound are similar to those found in games, and there are some cases where there are significant functions found in games that are not found in film sound.

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, discussed in the previous chapter. As shown, games are increasingly used as marketing tools and have become part of media franchises that may include film or television spin-offs. The interplay between audience and audio has other affects on the ways in which popular music is consumed. In the case of *Guitar Hero* or *SingStar*, there is a direct participatory and performance aspect to listening to the songs, as discussed in the previous chapter.

The direct participation between a player and the audio takes on a new role in rhythm-action or other kinetically based 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 (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 to which the player must respond.

As discussed in chapter 6, aside from rhythm-action games, licensed popular songs are most often found in sports games or racing games. These games are so popular that Electronic Arts created the brand name EA Sports to produce sports games, beginning in 1993, and is now the leader in the genre. EA Sports maintains some exclusive licenses with sports associations (Nascar, NFL, NCAA). The games aspire to high realism, as the motto for EA Sports would suggest: ''If it's in the game, it's in the game'' (later abbreviated to just ''It's in the game''). Racing games have also been very heavily reliant on popular songs, particularly dance and hip-hop games. *Need for Speed: Carbon* (EA Black Box, 2006), for instance, features Tiga, Goldfrapp, Lady Sovereign, and Roots Manuva, among others. To a certain degree, an interference or intertextual referencing between the games and the paramusical phenomena associated with these artists (videos, concerts, album covers, and so on) may occur, and indeed, may be intended in some cases (particularly connected are the urban, illicit aspects of certain games to hip-hop music, for instance). Genres such as racing games are much better suited to linear music than many other genres, since the player may be tied to a specific length of track or a particular length of time, and therefore, the timings are more predictable than in other genres. A racing track may last a minimum of three minutes, or a soccer half will last forty-five minutes, for instance, whereas a battle scene is far more temporally unpredictable. In addition, there is a well-established association of cars and motorcycles with popular music—particularly through hit films aimed at a younger audience, such as *Days of Thunder* (Paramount, 1990), *The Fast and the Furious* (Universal, 2001), or even older films such as *Easy Rider* (Columbia, 1969), which have set precedents for the use of popular songs in similarly themed games. Likewise, most popular sports teams have theme songs, and sporting games now are often multimedia ''events.'' In fact, one popular song used by many NHL arenas, Zombie Nation's ''Kernkraft 400,'' was based on a video game song (David Whittaker's music for *Lazy Jones* on the Commodore 64). The use of popular songs in sports-themed games seems fairly ''natural,'' then, in that it is a re-creation of these events.

One of the difficulties with using linear music in games, as noted in the previous chapter, is that it usually lacks the ability to adapt to the on-screen action, and therefore often fails to fulfill many of the other functions of games music. A crucial role of music and sound effects in games is the preparatory function that the audio serves, for instance, to alert the player to an upcoming event, or to forewarn the player of approaching enemies. Anticipating action is a critical part of being successful in many games, particularly adventure and action games. Notably, *acousmatic* sound—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, pp. 71, 85). This function, though present in films (even if we cannot force the camera to go in that direction, we mentally look there, argues theorist Michel 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, therefore affecting player decision making.[5] For instance, in *Ocarina*, 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 when traversing that particular area in order to decide the appropriate time to move. Stealth games make particular use of this function, as the player is alerted to enemy presence. Although it is possible to play games with audio turned off, this can considerably lengthen the learning curve of a game, and can make gameplay frustrating. Author Aaron Marks (2002, p. 190) 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'' (see also Jørgensen 2008).

Particularly important to games is the use of sound symbols to help identify goals and focus the player's perception on certain objects. Music may focus one's 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'' (Cohen 2001, p. 258). However, sound effects, such as footsteps or gunshots, more commonly serve this purpose. 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. In *Ocarina*, for instance, the lesser enemies all have the same or similar music, and beneficial items like gems or pieces of heart likewise all have the same or similar-sounding cues. The use of recurrent musical themes can also 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 his or her whereabouts in the narrative and in the game. In *Ocarina*, 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 he learns to play the theme on the ocarina, it sounds familiar, and when he must recall the theme at specific points in the game, it will be more easily recalled (Whalen 2004). Gerard Marino (2006) discussed his use of *leitmotif* in the game *God of War* (SCEA, 2005), as a way of matching the pace and emotion of the game:

> My initial task was to nail down the main theme of the game. I first focused on the main character, Kratos. The notes I had to go on were "the darkest, most brutal, most evil character ever," and I was given artwork and shown some gameplay. They wanted the music to reflect his soul-consuming quest for revenge on Ares, the Greek God of War, to help the player feel that Kratos will only be redeemed when he exacts his revenge. I rang up my Greek buddy Kostas and got him to translate some words and phrases for me, and settled on writing a theme that contained a rhythmic motive based on the Greek words "Ekthikisi" and "Litrosi" which translate to "Revenge" and "Redemption," so every time Kratos' theme plays his motivation is reinforced. . . . I initially multitrack recorded myself chanting/shouting/whispering the words 32 times to get a "choir-of-me" effect going at the opening and closing of the theme, which I later turned into an Intakt patch so I could drop the words/whispers into any piece I was writing and have it conform to the tempo in question, so this motive shows up often, especially in the cut scenes.

Music can also be used to enhance the overall structure of the game. This includes 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 the 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, p. 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 influence 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 personalize games, as has been explored in audio games designed specifically 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, p. 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, or timbre of their voice, while voice-over narrations can let us access a character's thoughts and feelings (ibid.). In *Grim Fandango*, the stereotyped accents and dialects 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 also 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 humor. Although much of the verbal interplay between player and game has traditionally been text based, with lines selected or typed in by the player, games are becoming more vocal, with players literally speaking to a game's characters.

Part of the role of dialogue—and audio in general—is the suspension of disbelief, adding realism and creating illusion. The illusion of being immersed in a three-dimensional atmosphere is greatly enhanced by the audio, particularly for newer games that may be developed in full surround sound, although 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 follows 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. This function of game audio does not differ significantly from that of film; but 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 him- or herself in such a massive setting, reducing confusion and frustration.

Another important immersive element Gorbman (1987) and Berg (1973) both discuss in relation to film is the historical function of covering the distracting noises of the projector in the era of silent movies. A similar function may be attributed to game 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 other sounds made by the surrounding environment (Cohen 1999, p. 41). Although perhaps to a lesser extent 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. Argues Oliver Grau (2003, p. 348) in his discussion of cinematic use of immersive techniques, such attempts at immersion are "part of endeavors to extend or overcome the constraints of the film screen." Surround sound technology is one way that audio is employed to conceal or reduce the "actual illusion medium by keeping it beneath the perceptive threshold of the observer" (ibid., p. 340). Audio helps to overcome the two-dimensionality of the image, to help the player feel immersed in a three-dimensional space, particularly in 3D games, as discussed in chapter 4.

Adding to the immersive effects of gameplay is the communication of emotional meaning, which occurs in game 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, cited in Cohen 2001, p. 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. In this way, sound works to control or manipulate the player's emotions, guiding responses to the game. Where games differ from linear media in terms of this relationship is that the player is actively involved in making decisions for his or her 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 immersive) relationship between the player and the character(s).[6]

IMMERSION AND THE CONSTRUCTION OF THE ''REAL''

Immersion, ''characterized by diminishing critical distance to what is shown and increasing emotional involvement in what is happening'' (Grau 2003, p. 13), is a subject of much debate within the industry and within academia. Salen and Zimmerman (2003, p. 450) argue that the immersive quality of a game comes not from the game itself, but through play, referring to what they term the ''immersive fallacy,'' which they define as ''the idea that the pleasure of a media experience lies in its ability to sensually transport the participant into an illusory, simulated reality. According to the immersive fallacy, this reality is so complete that ideally the frame falls away so that the player truly believes that he or she is part of an imaginary world.'' On the other hand, writer Andrew Glassner (2004) goes so far as to describe various levels or degrees of immersion within a game.[7] He begins with curiosity, or the casual desire to know. The next stage is sympathy, in which the player starts to see the world through the eyes of the protagonist. Once the player sees through the protagonist's eyes, the player can identify with the protagonist: seeing elements of the character in him- or herself, and elements of him- or herself in the character. From there it is possible to reach a state of empathy, or emotional bonding with the character. Finally comes a state of transportation, where the player can temporarily lose the boundary between him- or herself and the character (Glassner 2004, pp. 81–82). Ermi and Mäyrä (2005, pp. 7–8) argue for a three-part division of immersion: the first dimension is sensory immersion, in which ''large screens close to the player's face and powerful sounds easily overpower the sensory information coming from the real world, and the player becomes entirely focused on the game world and its stimuli.'' The second form is challenge-based interaction, in which ''one is able to achieve a satisfying balance of challenges and abilities.'' Third is imaginative immersion—what others discussed have meant when referring to immersion—in which ''the game offers the player a chance to use her imagination, empathise with the characters, or just enjoy the fantasy of the game.'' My focus here is on this imaginative immersive quality, which is strongly enhanced by audio.

The degrees of immersion experienced by a gamer are probably closest to what Oliver Grau (2003, p. 13) argues: ''obviously, there is not a simple relationship of 'either–or' between critical distance and immersion; the relations are multifaceted, closely intertwined, dialectical, in part contradictory, and certainly highly dependent on the disposition of the observer.'' Moreover, immersion may be a quality that comes and goes, depending on the mindset of the player. Regardless of whether or not immersion exists to any significant extent, it is a state to which most game developers aspire.[8] The construction of a believable, realistic space is a significant part of this drive toward immersion. John Belton describes the similar aspirations of realism in film: ''The direction of technological change and the development of sound practice answer, in part, the demand of classical Hollywood cinema for a means of illusionistic production that remains, for the most part, invisible—that is, it will not disturb the willing suspension of disbelief that permits audiences to become absorbed in a film's narrative or diegetic world'' (Belton 1999, p. 233).

Audio plays a significant role in the immersive quality of a game. Any kind of interruption in gameplay—from drops in frame rate playback or sluggish interface reactions—distracts the player and detracts from the immersion and from audio's playback—particularly interruptions in music such as hard cut transitions between cues (see Kline, Dyer-Witherford, and de Peuter 2004, p. 20). Many game audio devices are sold describing the more ''realistic'' experience the gamer will have upon purchasing these devices, such as the Nintendo DS headphones, Sound Blaster Live! soundcard, or Creative speakers, all of which promise to make games ''come alive.''[9]

Games are rarely set in realistic worlds, however, in the sense that they do not try to re-create present-day life on Earth. In many ways the realism aspired to in games is not a naturalistic realism in the sense of being a simulation of reality, but a *cinematic* realism that relies on established motion-picture convention. The ''cine-real'' is a sense of immersion and believability, or verisimilitude, within a fantasy world. It is the imagined real of the sound of explosions in space (which should technically be silent), or of the clarity of sounds underwater (which should technically be muddled). A realistic immersive space is developed with a ''naturalness ... mediated by different understandings of perception'' (Lastra 2000, p. 191). As James Lastra (2000, p. 207) discusses with regards to film sound, real sounds are not always the most appropriate sounds for film. Moreover, techniques of recording, as in film sound, have gravitated toward a specific aesthetic that is not necessarily the most ''natural'' sounding.[10] Sounds can be metaphoric and meaningful on levels other than just as an attempt to approximate reality, as sound designer Walter Murch (2000) argues: ''This metaphoric use of sound is one of the most flexible and productive means of opening up a conceptual gap into which the fertile imagination of the audience will reflexively rush, eager (even if unconsciously so) to complete circles that are only suggested, to answer questions that are only half-posed.'' Sound is as much an aesthetic choice as it is a reproduction of the imagined space, Murch argues:

> This reassociation of image and sound is the fundamental pillar upon which the creative use of sound rests, and without which it would collapse. Sometimes it is done simply for convenience (walking on cornstarch, for instance, happens to record as a better footstep-in-snow than snow itself); or for necessity (the window that Gary Cooper broke in *High Noon* was made not of real glass but of crystallized sheeted sugar, the boulder that chased Indiana Jones was made not of real stone but of plastic foam); or for reasons of morality (crushing a watermelon is ethically preferable to crushing a human head). In each case, our multi-million-year reflex of thinking of sound as a submissive causal shadow now works in the filmmaker's favor, and the audience is disposed to accept, within certain limits, these new juxtapositions as the truth. (Ibid.)

In games, as in film, the sounds themselves are largely constructed and assembled. But in games, reality in sound design is never an ''original'' recorded production-space sound, and rarely a raw recording of a ''real object,'' but usually a make-believe construction of sounds and synthesizer patches—a simulacrum of the real. For instance, if I need a monster roar for my game, I cannot use a production ''monster'' recording. I may take a gorilla's scream and a human recording of a yell, put them together, add some bass, add some overdrive, perhaps a little high-end scrape to sweeten it further. Even ''realistic'' recorded sounds such as gunshots are often treated with various effects to make them ''more real than real.'' It is not just sound effects that are simulated in games: the music may also be an unreal construction of multilayered simulacra. Troels Folmann has described his approach to creating realistic orchestral sounds in his music for *Tomb Raider: Legend* (Eidos, 2006), in which all instruments began as high-quality samples (in this case, primarily by the East West company, who advertise the ability to ''breathe life into the virtual world, creating a real space'').[11] From there, Folmann added ''random chaos''; the sound of the turning of pages of the orchestra, a quiet cough, the soft scrape of a chair, and the construction of a multitude of reverberation patterns to simulate the presence of a real orchestra. They may be too quietly placed in the mix for the conscious mind to detect, but, argues Folmann (2006), they help to trick our mind into believing we are hearing a real orchestra.

Sound director Rob Bridgett likens two growing aesthetic parallels with Hollywood and games in the mixing of IMAX and *ride films*, in which surround sound plays a vital role. The simulation ride film, or motion simulator, typically take place in custom-built rooms, and were developed in the mid-1980s but have continued to grow at an exponential rate. The main purpose of ride and IMAX films is a kind of sensory overload, an emotional assault, a complete immersion in the environment. Audio tends to be mixed very loudly and with a very physical effect due to significant use of subwoofers. In such films, audio is ''magnified and physically experienced way beyond reality . . . [they are] about entertainment, short, sharp and shocking. It is this where the subwoofer becomes the weapon of choice in the armory of the sound designer, to provide the necessary shock and awe for the audience, who generally are there to be 'wowed' '' (Bridgett 2005). Sound effects are placed very high in the mix and are designed to make full use of at least six channels, particularly the subwoofer. In this sense, the immersion can come as much from the physical *effects* of sound, as from their *affects*. This use of sound is somewhat determined by genre, as action shooters place more emphasis on the use of the subwoofer than, for example, a role-playing game: ''Shoot-em ups may play on either a tension and release narrative structure, one subtle build-up and other part all out kill-death frenzy, or they may opt for the full out continual derangement of a ride film, in which case prolonged immersion may become a problem for the player'' (ibid.). Similarly, immersion may be more of a goal in a first-person point of view game than in a ''god's view'' (overhead) simulation game, for instance.

CONCLUSION

Audio in games functions in a wide variety of ways, and, as shown, removing audio from games (turning the sound effects off, substituting the music, and so on) can significantly affect gameplay. This aspect, however, is dependent to some extent on genre and platform. For instance, Galloway discusses the nature of arcade games in relation to home games as follows: ''Arcade games are generally installed in public spaces and require payment to play: computer and console games, on the other hand, exist primarily in the home and are typically free to play once purchased. This material difference has tended to structure the narrative flow of games in two very different ways. Arcade games are often designed around the concept of lives, while console games are designed around health. . . . Arcade games are characterized by a more quantized set of penalties and limitations in play: one quarter equals a certain number of lives'' (Galloway 2006, p. 33). As such, the platform can set the length of games, which, as shown in the first three chapters, directly influences the length of music within a game. Mobile games need the ability to function without sound, as these are often carried into public spaces, whereas home games may need audio to distract from the environmental sounds in order to help immerse the player in the game.

Platform also directly affects genre. Arcade games today are nearly exclusively racing, first-person shooters, rhythm-action, or sports games. Such games are often predisposed toward more linear music, and therefore toward licensed music, and they also have very specific identifiable target markets who venture out to public spaces such as movie theaters where arcade games are commonly now housed.

Genre also affects the style and rules of gameplay. Many role-playing adventure or *Sims*-style games, as well as online games, have extended lifetimes, where players typically save and come back to the same game over several days, weeks, or even months. The use of sonic symbols is particularly necessary in these games to provide a sense of cohesiveness, and to help guide the player along in terms of the narrative and his or her location in the game matrix. For the composer, these lengthy games require much more variation in the soundtrack, since the player will be engaged with the sound for a much more significant amount of time—a difficult problem that can be overcome in a variety of ways, taken up in chapter 8.