Philos. Technol. (2014) 27:235–249
DOI 10.1007/s13347-013-0113-z

SPECIAL ISSUE

Musical Ecologies in Video Games

Michiel Kamp

Received: 1 November 2012 / Accepted: 10 June 2013 / Published online: 26 June 2013
© Springer Science+Business Media Dordrecht 2013

Abstract What makes video games unique as an audiovisual medium is not just that they are interactive, but that this interactivity is rule bound and goal oriented. This means that player experience, including experience of the music, is somehow shaped or structured by these characteristics. Because of its emphasis on action in perception, James Gibson's ecological approach to psychology—particularly his concept of affordances—is well suited to theorise the role of music in player experience. In a game, players perceive the environment and gameplay situations in terms of the goal-oriented actions they afford. Nondiegetic music, not tied to any place in the digital game world, can play a unique role in the structuring of these affordances. Through a series of case studies, I will show that music creates and structures situations both in the game environment (such as the appearance of enemies in Unreal) and beyond the game environment (such as the death of an avatar and the restart of a level in Super Mario Bros.).

Keywords Ecological psychology · Music · Affordance · Diegesis · Video games

When we hear music in video games, what do we hear, and why do we hear it? It is certainly unlike typical forms of musical listening, such as the kind of listening we do in the concert hall, or to a CD at home. But in many ways, it is different from hearing film music as well. In a sense, most listening to video game music is much more like normal, everyday listening, in that it is action oriented (Gaver 1993). Unlike in musical or aesthetic listening, we are normally more concerned with the source of the sounds we hear, and how it relates to our situation, than with the sounds themselves. Upon hearing the sound of a car when trying to cross a street, we do not stop to admire the slow crescendo, but identify it as approaching us, gauging its direction and its speed. When pouring tea, we hear a small change in timbre in the sound because it indicates the filling of the mug, not merely because it pleases us. In

M. Kamp (✉)
University of Cambridge, Cambridge, UK
e-mail: mk646@cam.ac.uk

short, what we hear and how we listen is determined by our goals and intentions, and our relation to our environment: the busy commuter will attend to the soundscape of a city differently than the leisurely shopper. When playing video games, we hear in much the same way—oriented towards actions, specifically rule-bound actions—and this article will concern itself with the question of how background music in games fits into such a way of listening.

One theory of perception that concerns itself with action is James Gibson's ecological approach to psychology. I want to adopt some of his ideas—specifically those of the relationship of an organism to its environment, and of perception in terms of *affordances*—to offer an account of hearing music in games that is based in player (inter)action. After introducing Gibson's approach and the way it has subsequently been taken up by researchers both in musicology and video game studies, I will move on to a series of case studies that illustrate the unique place of background music in a game's ecology, and how music can be heard in this way across multiple game genres. The first-person shooter *Unreal* will show how music is perceived as structuring situations and their affordances. Then, by looking at how music functions around the death of the player's avatar in *Super Mario Bros.*, I will suggest how music can structure a game's ecology beyond its diegesis. Finally, I will consider how a musical ecology would work in an abstract game such as *Tetris*.

1 Gibson's Ecological Psychology

Gibson developed his ecological approach to the psychology of perception in a series of books in the mid-twentieth century. Influenced by the work of pragmatist philosopher William James and the *Gestalt* psychologists, Gibson's ideas ran counter to the predominant cognitive psychology of the time, which focused on discovering the mental processes and systems that process sensory information from the outside world as if it provided a "kaleidoscopic inflow of sensations" (Gibson 1966, p. 5).[1] Gibson argued that organisms need not decode complicated and chaotic information from the outside world: they have a sensory system that *resonates* with their environment (ibid., p. 267). This resonance follows from process of mutual "tuning", in which the environment shapes organisms to perceive and react to it in a certain way, just as the organism shapes and influences it in turn. Thus, an animal's perceptual system evolve in tandem with their environment to extract certain information necessary for the attainment of their goals, which include things like finding food and shelter. Similarly, throughout pedestrians' lifetimes their hearing is attuned to the city to listen for oncoming traffic as they cross a busy street.

Considering the perceptual system in terms of a relationship between organism and environment leads Gibson to two important concepts: *invariants* and *affordances*. Both concepts stress the importance of action for perception. The first, invariants, are features in the environment that remain persistent to our senses as we move through it. Our perception is structured to register and take into account these invariants (Gibson 1979, p. 61). Some examples of visual invariants that Gibson mentions are motion parallax (ibid., p. 76) and the density of optical texture gradients (ibid., p. 67).

[1] Or, in William James' words, the world as "one great blooming buzzing confusion" (James 2007, p. 488).

Affordances (ibid., p. 18) are possibilities for (inter)action—specific patterns or relationships—between organism and features of the environment. The kinds of actions that an object in the environment affords are dependent on the morphology of the perceiving organism, their goals and the environment. Thus, the same object can afford different things to different animals and to the same kind of animals in different situations. Musicologist Eric Clarke gives the example of a termite's perception, which has been tuned to a wooden chair so that it affords eating. For a human, this chair, in normal circumstances, affords sitting; for a person under attack, however, the chair affords self-defence (Clarke 2005, pp. 37–38).

Gibson, working in the psychological tradition of William James, thought his ecological approach should ultimately be empirical in nature, and phenomenological description was no more than "a good starting point for identifying significant psychological phenomena that warrant more detailed analysis" (Heft 2001, p. 114). Nevertheless, many of Gibson's own examples of invariants and affordances are much like phenomenological descriptions, and there are striking similarities between the observations of some later ecological psychologists and phenomenologists like Don Ihde.[2] Dreyfus and Kelly, in their phenomenology of experience, even use the concept of affordances explicitly to describe a kind of knowledge we have about the world that is not a belief. Rather, it is a kind of pre-reflexive, immediate "sense" we have of objects in the environment (2007, p. 52). They admit that their phenomenological take on the concept is at odds with Gibson's, who explicitly argued that affordances are objective relationships in the environment, not just categories of experience. My own application of the ecological approach will follow this more phenomenological tradition, since it has been taken up in similar ways in musicology and game studies. Nonetheless, the idea that an affordance is an objective relationship between organism and environment is important to my conception of video game ecologies, as I will explain in the next sections.

2 Ecological Approaches to Music and Video Games

Both musicology and video game studies have adopted Gibson's ecological approach in their own way, with their own agendas. Musicologists like Windsor (2004); Windsor and de Bézenac (2012) and Clarke (2005) are interested in applying it to Western classical music to explain (away) the idea of autonomous art music.[3] For them, traditional musical listening modes, such as attentive listening in the concert hall (or at home on the couch to a recording), are just as action-based as our normal, everyday listening. Clarke argues that "there is really no such thing as passive listening" (2005, p. 205): motion in music is an invariant and its perception is direct (ibid., p. 89), and moreover music *affords* movement, whether we act upon it or not.[4] Windsor addresses the problem in a similar way to Clarke, by arguing that "perceived relationships between musical motion and real motion and real movement are not

[2] Compare Windsor and de Bézenac's examples of intermodal variables (2012, p. 105) with Ihde's on detecting object shapes through sounds (2007, pp. 62–63).

[3] For a more complete overview of the ecological approach to music, see Windsor and de Bézenac 2012.

[4] In this regard, his arguments are similar to those from the field of embodied and situated cognition (e.g. Johnson 2007 and Leman 2008).

arbitrary or metaphorical, but motivated by integration of action and perception" (Windsor and de Bézenac 2012, p. 112), backing his claims with empirical evidence from cognitive neuroscience. The question is how their ecological account of concert hall listening can be applied to hearing music in multimedia contexts such as video games. Krueger offers a way by pointing out the "joint-attention" that takes place during musical listening at live performances (2011, p. 17). While for Krueger this kind of listening is still focused primarily on the music, our awareness of other people listening for the same things acts as "a kind of phenomenal filter or framework through which the explicit objects of attention (e.g. the musical piece) are given" (ibid., p. 17n). I want to argue that in the case of background music in a video game this "phenomenal filter" is inverted, so that the music becomes the filter through which we experience the game. The question then is *how* background music filters our experience, or more appropriately, what it affords in those situations, and to what extent these are the same affordances as those that Clarke et al. describe.

In video games, one can imagine that ecological psychology's action-based account of perception would naturally fit quite well, but its applicability runs deeper than that. One could say that each video game represents its own unique ecology, where the organism and environment are not shaped by millions of years of evolution, but designed by game developers and hardware engineers; environments in the form of levels and stages in games, organisms in the form of avatars and control schemes that represent the player. This leads to two conceptions of video game ecologies: one in which affordances are hard coded into the game, and one "softer" account in which they depend on player perception, effectively mirroring Gibson's physical account and the more phenomenological one. An example of *hard* affordances comes from Pinchbeck (2009), who categorises objects in a number of First-Person Shooters (FPSs) in terms of their coded possibilities of interactions, which he identifies as affordances. A crate, for instance, can be designed to be climbable, pushable, breakable, openable, etc., which leads to other properties that have to be accounted for such as whole vs. broken. Linderoth and Bennerstedt adopt the *soft* account in their discussion of player learning in games, which for them is "a process of becoming attuned to certain aspects of the environment in such as a way that we gain new affordances, new ways to act and interact with the world" (2007, p. 602). This is very much related to Goffman's concept of the "rules of irrelevance" that Juul (2005) uses to describe player learning in video games: as we devise the right strategies to play the game, we learn to discern the cosmetic, "fiction" elements from the rules and goals of the game. For Linderoth and Bennerstedt, new affordances can arise as players learn more about a game, as situations and our understanding change.

The primary question I want to address in this article is "how does music fit in a video game ecology?" Unlike objects or sounds, the ecological status of the musical soundtrack in a video game is less obvious. Can it be said that the music affords actions in the same way as the crates in Pinchbeck's FPSs do, or does it aid in attuning the player to a game environment, following Linderoth and Bennerstedt? The question is further complicated by the notion of diegesis in video games. A crate can easily be said to be a diegetic object, whereas most musical soundtracks are nondiegetic—not part of the game world, and therefore perhaps not part of the environment either. Yet, unlike films, from which the concept diegesis stems, a lot of video games do not let themselves be subjected so easily to a diegetic/nondiegetic

binary; after all, what is diegetic in *Tetris*? The blocks? I will start my discussion with the work of Mark Grimshaw who in his PhD dissertation (2007) and its subsequent publication as *The Acoustic Ecology of the First Person Shooter* (2008) adopts Gibson's ecological approach to perception. Through a series of case studies, I will formulate a thesis of how music fits in such an ecology. Then, I will discuss the "twofoldness" of ecological perception that music exposes in the FPS genre, and discuss the ecology of nondiegetic music in other genres to contrast this.

3 Diegesis, Sound and Music

Before moving on to how Grimshaw constructs his acoustic ecology, I want to briefly explain the concept of diegesis, as it is implied throughout many of his arguments, and will prove to be important for how we perceive a musical ecology in many games as well. The concepts *diegetic* and *nondiegetic* have been traditionally used in film music scholarship to denote the difference between music that originates in a film's narrative world (or diegesis), and music that does not (Cooke 2008, p. 9). The term has been taken up in video game scholarship as well, but this has led to some complications that fall into two categories. The first has to do with avatar–player interaction and has been addressed by Jørgensen (2008) in particular; the second revolves around Juul's idea that video games can have incomplete and incoherent fictional worlds (Juul 2005, p. 123), and this is related to the problem of the *Tetris* blocks I mentioned above as well.

Interactivity complicates the notion of diegesis by creating situations that cannot be fully explained in the story world. When the player's avatar in a video game suddenly blurts out "I can't do that," or "I'm low on mana," who are they talking to? The answer of course is "the player," but their lips are moving, and their voice *sounds* like it is coming from the game world—it's not a voice over. Jørgensen calls this kind avatar–player interaction *transdiegetic*. Some musical effects can be labelled as transdiegetic too. Imagine a diegetic scenario wherein a character is walking down a hallway and suddenly draws their weapon. Seconds later, a monster appears around the corner. What prompted their sudden vigilance? A sixth sense? The experience of a seasoned adventurer? Again, the answer is of course "the player," who heard nondiegetic music signifying danger. This too can be construed as a case of transdiegesis. However, I am concerned with how the player *hears* this music as they are interacting with the game. Unlike diegetic sound effects, background music does not "move" between audio channels or increase or decrease in volume to correspond with movement through the game world (in Grimshaw's words, it does not create "resonating spaces", as discussed below). Nor does this kind of music have a visualised sound source, such as a record player, or a performer and their instrument.[5] All these factors contribute to the fact that the player does not hear the music as a phenomenon that is "out there", to be looked for in the diegetic environment. While its transdiegetic qualities pose a theoretical problem, for a player playing a game,

[5] Grimshaw (2012, p. 350) also suggests that "the design of many digital games (especially FPS games) does recognise a distinction between diegetic and nondiegetic sound through the inclusion of separate volume controls for sound effects and music."

background music is phenomenologically nondiegetic.[6] In this nondiegetic role, music can function like the "phenomenal filter" Krueger describes (see above), altering the player's experience of objects they attend to while remaining "invisible" (see Farrow and Iacovides 2012).

The second problem with diegesis in video games is that, depending on the genre, their fiction and consequently their diegeses are not always consistent or coherent. Juul asks the question "Why does Mario have three lives?" (Juul 2005, p. 123). Why not two or four? How can his deaths, retries, and subsequent "game over screen" be explained by the game's fiction? Does he climb back out of the pits he falls in, but is forced to give up after three times? Is he reincarnated three times? Juul's point is that because *Super Mario Bros.* (Nintendo 1985, henceforth *Mario*) is a game which has rules to "fill in the gaps" the fiction leaves, it does not need to be explained. For all intents and purposes, *Mario* has multiple diegeses, each one spanning one "life" of Mario. This leaves room for nondiegetic elements like the interface and the music to fulfil unique roles that surround and relate those diegeses to each other in terms of what they afford and how, as I'll try to argue further in my case study of *Mario*. Then there are abstract games like *Tetris*, which need not have a notion of diegesis at all. Their ecologies are constructed and perceived in a different way from representational games, and I will briefly discuss them in a separate section after my main argument. For now, I will restrict my discussion to games with a diegetic game world which represents real-world environments and their affordances, and I will adopt a simple two-plane notion of diegetic/nondiegetic, where music is not presumed to be part of the game world.

4 A Musical Ecology of First-Person Shooters

Sound is of particular importance in FPSs. The genre offers a relatively limited field of vision—as the player is looking through their avatar's eyes—but often requires a high grade of situational awareness. Unlike other genres, which can rely on bird's eye camera views or mini maps to indicate the location and distance of objects, FPSs rely to a large extent on audio. Moreover, games in the genre often insist on a high level of realism, which requires a particular fidelity of graphics, but also sounds. For Grimshaw, the acoustic ecology of an FPS is created through the interaction between a player and a soundscape (Grimshaw 2007, p. iii). Apart from affordances, it comprises several acoustic "spaces", divided into *resonating spaces* and *paraspaces*. A resonating space defines an avatar's location, and the size and shape of their environment, and is created through player interaction with sound effects that move from channel to channel (left to right for instance) based on their avatar's movements (ibid., pp. 176–177). A paraspace is created by sounds that locate the player in a certain time and *place* through cultural connotations (ibid., p. 197). Player interaction with a single sound effect can aid in the creation of multiple spaces: the sound of gunfire can create a post-industrial paraspace, and a resonating space that defines a gun being fired somewhere close, to the left of the player's avatar.

[6] See also Grimshaw, Tan and Lipscomb (forthcoming) for a short discussion on the differences between the notion of diegetic/nondiegetic sound in films and video games.

Grimshaw explicitly states that "only sounds that derive from game world actions and events … form part of an understanding of [a] game's acoustic ecology" and that "any sounds heard while playing the game that do not inform about the game world (such as the musical score …) are not considered a part of the acoustic ecology framework" (ibid., p. 221). These statements contain two separate charges against music's claim to be part of an acoustic ecology. The first is that only diegetic sounds contribute to its creation, the second that music does not inform about the game world, or diegesis. The second charge can be easily dismissed. Like the sound of gunfire, music's cultural connotations can aid in the construction of paraspaces, for example by using exotic scales or instrumentation. The first charge is less specific and more difficult to answer. I'd like to begin by looking more closely at how a FPS's ecology—and its affordances in particular—are constructed.

Whether or not it's part of an acoustic ecology, nondiegetic music is a part of the soundscape of almost every FPS to some extent. Classic examples of the genre like the *Half-Life, Call of Duty*, *Unreal* and *Halo* series all feature musical soundtracks. So how does music work in these games? An important use that returns in many FPSs is so-called "danger state music" (Whalen 2007) to indicate the presence and absence of enemies and other dangerous situations. When music starts playing, players know they should be more alert, and that danger could be lurking while it lasts, even though there is no visual indication. But in many FPSs the use of music to indicate gameplay "states" can be more intricate than this. As an example, I want to discuss three musical events and the relationship between them in the first moments of *Unreal* (Epic Games 1998).

In the first one, right at the beginning of the game, the player's nameless avatar starts on what appears to be a prison spaceship, where something has gone horribly wrong. The cell doors are open, red alarm lights are flashing, and in the distance screams can be heard. There is no music at the very start of the game, but as the player makes their way to a control room a voice can be heard saying "you have entered a restricted area," after which cue A starts. It consists of distant metallic drums—reminiscent of the soundtrack to the *Terminator* films—without any recognisable rhythm or beat, which after some time is supplemented with synth-pad harmonies. When the player once more hears the voice saying "you have entered a restricted area," the voice recording suddenly breaks down and with it the music fades out. This is immediately followed by the player approaching a door, behind which a short "auditory cutscene" takes place.[7] The sounds of screaming (coming from both a human and a monstrous voice) and gunfire indicate a fight, and after the door opens the player finds a corpse and a pistol. When they pick up the pistol—their first weapon in the game, an iconic occasion in FPSs—cue A abruptly starts again.

In the second event, the player is exploring the outdoor environment directly outside the crashed prison ship. The musical cue that is playing during this, cue B, has similar instrumentation to cue A, but more consonant harmonies and a light mid-tempo beat that fades in and out as the player is exploring the terrain. When the player

[7] A normal cutscene is a short sequence containing some narrative exposition during which the player is unable to interact with the game. The trick *Unreal* uses to achieve similar narrative progression is to momentarily halt the player's progress by locking the doors to the room they are in, and having sounds come from the next room.

enters a tunnel they encounter the first enemies of the game, two brutes, which attack them on sight. Cue B crossfades into cue C, an archetypal example of danger state music that features a heavy industrial metal beat, with a fast, distorted ostinato that mimics an electric guitar riff, every now and then interspersed with synth-pad harmonies to maintain thematic unity with cue A and B. When the player kills the brutes, cue C does not fade out, but plays until the player reaches the end of the tunnel, where the player encounters a third and final brute. While traversing the tunnel, the music lingers for several reasons. First, it marks the gravity of the player's first combat encounter, and can be said to represent the adrenaline rush. Second, it keeps tension for slightly longer, to keep the player on their toes. And third, it equates not just the dangerous situation with the presence of brutes, but also with being in a tunnel. When the player finally exits the tunnel, cue C crossfades back into cue B.

As the player enters an underground area slightly later in the game, the game loads a new map. Cue D starts abruptly as soon as the map loads. Its immediate and continuing presence, rather than the slow fading in and out of cues B and C, together with its clear beat and mid to high tempo are reminiscent of the omnipresent musical soundtrack of the original *Doom* (id Software 1993), one of the first FPSs, which featured nondynamic[8] music that spanned across an entire level. The similarity between *Unreal* and *Doom* is further emphasised when the player encounters more brutes and there is no dedicated danger state music, just cue D. Through its relationship to other music both inside the game (the danger/safety state music of cues B and C) and outside the game (the similarity to the soundtrack of *Doom*), the music has recontextualised the situation.

These examples show that players, through relating musical cues both to other cues within a game and to music in other games, assess their situation at least partially by relying on music. While the sound of an approaching enemy can create—in Grimshaw's words—a resonating space that affords turning towards an enemy and shooting it, music creates the *situation* whereby a player is more likely to interpret that sound as affording those actions. The encounters with the brutes when cues C and D play are different situations. The first brutes afford a cautious strategy: backing away from them while firing in their direction for instance. The sudden change in music indicates an unfamiliar situation, in which aggressive action is required, but the nature of those actions is still unknown. The continuous presence of cue D in the second situation means the player approaches the brutes more assertively, with a pace that need not change according to the music's suggestions.

So if music is important in how we perceive situations in a game, how can it fit into Grimshaw's account of an FPS's ecology? Let's revisit Grimshaw's second charge, points (1) to (3):

(0) An ecology consists of an organism and its relation to an environment;
(1) An acoustic environment is created through resonating spaces;
(2) Nondiegetic music does not create resonating spaces;
(3) Nondiegetic music is therefore not part of the environment.

[8] Whenever I use the term "dynamic" or "nondynamic", I'm not referring to dynamics in music terminology, but concepts introduced by Karen Collins (2008) to denote the different ways in which music is designed and composed to (cor)respond to player interaction. The more dynamic the music, the more it will change based on what the player does in a game: where they go, what enemies they encounter and so on.

But according to Gibson's definition of ecologies, (0), music can still be part of an ecology if it is part of the organism, and this is where I think we should locate music in Grimshaw's theory. This is possible when we think of the "organism" not as the player, but as a representation of the player in the game world, i.e. their avatar. Phenomenologically, the player's avatar occupies a unique role in the diegesis. While it is partially comprised of elements that the player *can* see and hear (such as the hands holding a gun at the bottom of the screen in an FPS), gameplay usually dictates that we see past those elements, but *through* them at the same time. This is a different way of "seeing through" from the one Aarseth (2004) has in mind when he says that "when I play [*Tomb Raider*], I don't even see [Lara's] body, but see through it and past it." It is more like Merleau-Ponty's description of how a woman with a feather in her hat will instinctively keep a safe distance between the hat and things which might break it off: "[s]he feels where the feather is just as we feel where our hand is" (Merleau-Ponty 1962, p. 165).[9] Following Farrow and Iacovides (2012), I want to extend the notion of an avatar to nondiegetic elements of the game that we perceive through as well, such as the viewing angle, control layout, physical controller, interface and nondiegetic music. The organism in a video game ecology is the player's avatar in its broadest sense; it is the result of what Calleja (2011, p. 169) calls the incorporation of the player into the environment, which at the same time results in the phenomenological incorporation of the environment in the player's perception.

5 Situations and Temporal/Event Structure

To summarise, from my case study of *Unreal*, I conclude three things about the perception of music in FPSs: (1) in the ecology of FPSs, music is part of the organism rather than the environment; (2) the player perceives the diegetic game world through their avatars, which are constructed partially through diegetic elements, but also through nondiegetic elements like the music; (3) music structures affordances by creating situations through reference to earlier situations in the game, and through connotations outside the game (i.e. paraspaces). This means that at least in the FPS genre, music is instrumental to the creation of situations which define affordances, and, by relating situations to each other, plays an important part in creating the game's perceived *structure*. So how do we adopt the ecological approach to other game genres? My emphasis on the issue of diegesis has caused me to relate music to the organism in FPSs ecology, but what is the "organism" in for instance a platform game? We do not literally perceive the world through Mario's eyes in the same way as we do in *Unreal*. We can say that, in Juul's words, the fiction of an FPS is much more coherent and consistent than that of other genres, like platform games. In the next section, I will discuss how music in other games' ecologies "deals" with these fictional inconsistencies, and to what extent the definition of music as prosthesis applies.

[9] In this sense, the contrast between Lara's "butch (her guns/athletic prowess)" and "femme (exaggerated breast size, tiny waist, large eyes, large mouth)" modes of representation that Helen Kennedy (2002) notes correspond roughly to the difference between the features we see *through*, i.e. her abilities, and the features we look *at*.

6 Structuring *Super Mario Bros.* Beyond the Diegesis

In many games, we perceive the environment and what it affords *through* our avatars and their abilities. In *Super Mario Bros.*, Mario's different sizes offer different routes through the environment. A large Mario can break blocks with his head, opening alternative routes; a small Mario is able to traverse narrower passageways to get to the end of the level. And on a more ineffable level, we incorporate the game's many pits and chasms in our ecology through our judgments on whether and how they afford jumping over, using our intimate knowledge of Mario's movements. But what happens when we judge wrongly, and Mario falls and "dies"? Is this still part of the game's ecology? How can we say it affords anything? What happens in effect is that one diegesis "ends" and another, unrelated one begins—although we take away expectations and experience from what happened before Mario died, we do not relate any of the *fictional* occurrences of one diegesis to the other. It is here, "between" these two diegeses, that—more so than in the coherent fictions of FPSs—nondiegetic music comes in to create a game's structure for the player.

The music in *Mario* does this very explicitly. The first thing the player notices is the complete silence in the menu screen. When the player starts a new game and is presented with their avatar Mario in the diegetic game world, the music begins playing, starting with a short introduction. The music—the "Overworld Theme"—then proceeds to loop through a series of melodies until the player finishes the level or Mario dies, at which point a coda is played, and the music stops abruptly.[10] The different diegeses that *Super Mario Bros.* presents to the player are explicitly underlined musically through these abrupt breaks, pauses and codas. What this offers or affords to the player is not just a moment of respite after failure, but first of all a hindrance in their progression through the game, as the player has to wait around while the music plays, falls silent and the level restarts. Through aligning the musical structure of introduction-loop-cadence with Mario's "lifespan", the music makes the repetition of diegeses stand out to the player.

We can speak of the music in terms of affordances if we look beyond the ecology of the diegetic game environment to the game as a whole. Whereas a pit might afford jumping over to the player *through* their avatar Mario, music structures the affordances of dying directly for the player. And it is able to do so specifically because of its nondiegetic status: it can stretch over diegeses, or intercut them with introductions and cadences, like in *Super Mario Bros*. But how does death as an affordance fit into a video game ecology? In the case study of *Unreal*, I took the environment to be the diegetic game world and the organism the avatar, both as it is represented to the player and as it incorporates the player in the environment through nondiegetic interface elements and background music. In *Mario*, I focused more on the other, "nondiegetic plane," where the environment is the game in its totality and the organism is the player. To the player, the death of *Mario* is merely an obstacle, affording the repetition of a level and a hindrance in one's progress. This also means that an "environment" does not necessarily need to be a spatial environment that in some way represents real-world spatial environments. It can be the game in itself, as a process in time with which the player

[10] The "Overworld Theme" is temporarily interrupted by other themes when Mario enters another area, or picks up a star powerup.

interacts. In the next section, I will briefly discuss how music structures affordances in games that completely forgo representation.

7 Abstract Ecologies

Most games feature at least some rudimentary form of representation, or fiction in Juul's terminology, but this fiction can have varying degrees of consistency and coherency. The more inconsistent or incoherent, the more abstract a game is. But there are games that forgo any kind of fiction, in which any representational elements are merely ornamental:[11] the cartoon animals in the background of *Peggle* (PopCap 2007) combined with the *Ode to Joy* chorus from Beethoven's Ninth are in no way to be interpreted as forming a coherent story or fictional world. Nor do these games have any clear distinction between diegetic and nondiegetic elements. How does one approach the experience of these games from an ecological perspective then? Here, there is no diegetic "plane" with a visible and audible avatar as organism and a game world as environment, but one can still take the perspective I took with *Mario* above, with the player as organism and the game in its entirety as environment.

The classic Game Boy version of *Tetris* (Bullet Proof 1989), arguably the most famous abstract game, actually has a spatial environment in the form of a playing field, which is clearly separated from the interface. The playing field, in conjunction with the game's controls, affords rotating the blocks to make them fit the rows at the bottom. Kirsh and Maglio (1994) have argued that these kinds of "epistemic actions" facilitate spatial cognition in *Tetris*, making it easier for players to complete the game—in ecological terms this suggests that the environment affords puzzle solving. The question is to what extent music can play these roles, or what other types of actions it can be said to afford. The music in *Tetris* (the Game Boy version at least) is in Collins' (2008) terminology nondynamic: before the game starts the player picks one of three pieces of music which will loop during gameplay until the player fails, after which a short game-over cadence plays in a very similar way to *Super Mario Bros*. The pieces include two arrangements (of the Russian folk song "Korobeiniki" and a minuet from a French suite by JS Bach) and one original composition, all three lasting around a minute. In spite of its similarities, the music of *Tetris* does not structure the event of failure and restart in the same way as *Mario*. This is partially to do with the game's abstract nature, as there isn't a "story" that is being told that the music is aligned with through an introduction-loop-cadence structure. Mostly it has to do with the difference in rules and goals between *Mario* and *Tetris* though: while in *Mario* there is a clear progression towards the end of a level that can be interrupted and delineated by the music, in *Tetris* the aim is to keep the game going for as long as possible—*Mario* is like an obstacle course, *Tetris* like plate spinning. But while *Tetris*' music might not create the same narrative "arc" (of variable length) as *Mario*'s, it can still be said to structure the playing experience. There isn't so much an arc that the music can span, but a continuous tension that can be broken at any point in time. The short, repetitive loops of the *Tetris* music add to this by affording a certain indeterminacy in the player's experience of the time played so far.

[11] This is quite similar to the use of fleurs de lys and leaves as ornaments or parts of arabesques in architecture, which Scruton (1997, p. 127) compares to musical imitation.

My example shows that in abstract games music can structure affordances just as well as in games with a diegetic game world and a narrative, even with a very simple nondynamic soundtrack like in *Tetris*. As Nattiez argues, music can "imitate the semblance of a narration without ever knowing the content of the discourse" (1990, p. 257), and, continuing on from his point, one can see that music can accompany a series of abstract events and still have this quality. Instead of a narrative arc, what is created by the music is a "gameplay" arc. In *Mario*, this arc stretched over multiple diegeses, but in *Tetris*, it is wholly independent of any notion of diegesis or narrative in a traditional sense. Considered in this way, music structures the gameplay of *Tetris* in time just like the music in *Mario* and *Unreal*.

8 Ecologies of Music Games

Until now I have exclusively considered examples of music as a background, the aforementioned "phenomenological filter" through which we experience the game environment. The ecological approach can also be extended to games that foreground music more, however. On the one hand, these games have the potential for a much wider variety of musical affordances than the dynamic transitions in *Mario* and *Unreal* or the nondynamic repetition in *Tetris* on the one hand, and on the other can even afford a new kind of hearing and even "making" music. A close integration of music and gameplay has been employed in a number of games in recent years. *Amplitude* (Harmonix 2003) and *Audiosurf* (ASCARON 2008), for instance, require players to manoeuvre at significant points in the music, such as on rhythmic or melodic accents. In these games, the music very literally structures the environment, as it determines where the obstacles are placed in time and space. *Audiosurf* allows the player to import a track from their audio library (in mp3 format for instance) and creates an obstacle course out of it, using the music's dynamics to generate blocks and elevation. While the course can be traversed successfully using purely visual information, knowledge of and listening to the music affords a better understanding of where and how to act. This happens on a much more intimate level than in *Unreal* or *Mario*, where music structures player perception of situations that span over a much larger timescale. In *Audiosurf*, knowing where and how long the more intense passages in a piece of music are affords a heightened attention at those points in the course, but on a smaller timescale, anticipating a downbeat in the music can even afford picking up a single block.

In many ways, the games in the *Guitar Hero* (Harmonix; Neversoft) and *Rock Band* (Harmonix) franchises are very similar to *Audiosurf* and *Amplitude*. Rather than manoeuvring, they require of the player button presses on the guitar-shaped controller at significant points in the music. And like *Audiosurf* and *Amplitude*, the music affords correct timing and a better understanding of the visual environment—in this case the shapes of the coloured dots on the highway—through what Shultz describes as "map[ing] musical time directly to physical space" (Shultz 2008, pp. 180–181). But this close relationship between music and visuals, no doubt aided by the player's hand movements on the guitar controller that mimic chord shapes and strumming on a real guitar, raises the question what is affording what in these games. As players engage attentively with the songs in *Guitar Hero* and *Rock Band*, the coloured

buttons on screen come to afford hearing certain melodic contours and structures in the guitar parts. Even apart from the question whether players are actively making music in these games (for more on this see Miller 2012), they afford new ways of hearing and understanding music.

A possible example of "making" music in a game is the downloadable puzzle game *Chime* (Zoë Mode Brighton 2010), which was inspired by *Tetris*. The game shares with *Tetris* its control style and gradual intensifying of the action and pacing, but the different goals change the game's structure and the way the music is perceived drastically. Each of the game's five levels is named after the track that its musical soundtrack consists of: *Brazil* by Philip Glass, *For Silence* by Paul Hartnoll, Moby's *Ooh Yeah*, *Spilled Cranberries* by Markus Schultz and *Disco Ghosts* by Fred Deakin a.k.a. Lemon Jelly. For the game's soundtrack, these tracks are broken down into layers of instrumental accompaniment, and their melody lines are "chopped up" into small runs and flourishes. As the player progresses through a level and starts filling up the grid, the layers of accompaniment are gradually brought in, and when the player places blocks in the correct positions the melodic flourishes are played randomly. This gives the sense that the track is slowly being constructed, which cross-domain maps onto the player's progression. As the name implies, *Chime* is presented as a musical puzzle game, and to some extent it can be said that the game affords music making, but in an unconventional way. On the one hand, as the music is a by-product of the player's progression through the game and the game's goals do not require direct interaction with the music (such as "playing" the melody by placing blocks in particular ways), music does retreat into the background of the game's ecology and perceptual environment. On the other hand, the close mapping of gameplay onto music means the music does not so much afford playing the game as playing the game affords making the music.

9 Conclusion

Adopting concepts from Gibson's ecological approach helps answer questions such as why the attention of players is drawn to the music at some points in the game and not at others, and how players use music to make sense of a game and the challenges and situations it poses. The case studies show how in a video game ecology, background music primarily constructs and structures affordances through its (lack of) dynamic transitions. Situations in *Unreal* arise when the game's code prompts a new musical cue, and death in *Super Mario Brothers* is musically structured by the dynamically inserted codas and Game Over cues. *Tetris*, on the other hand, shows how the music's lack of transitions and repetitiveness play a more important role. In abstract games and games with less coherent diegeses than FPSs, such as *Tetris* and *Mario*, the way music structures the gameplay and its affordances provides a useful way of looking at its perceived function. One can for instance compare the way music structures death in other platform games, like the continuous musical flow that affords the fast gameplay of *Super Meat Boy* (Team Meat 2010). Music games offer musical affordances in much higher quantities, structuring not just situations, but moment-to-moment actions as well, to the point where the player's actions can be said to afford new ways of understanding the musical soundtrack.

Of course, what I have outlined is a theory of *functional* listening to music in video games: we listen to music in terms of how it helps us achieve a game's goals. It is not a direct application of Gibson's psychology of perception, but rather a suggestion to think of video game listening in terms of some of its concepts. Part of playing a video game means becoming attuned to its environment—whether this is a three-dimensional simulation of a real-world environment, or something more abstract—and developing affordances along the way. That does not mean players cannot stop to enjoy the music for its own merits, or for how it reflects the particular visuals or setting the player is faced with. An account of these other ways of listening could proceed from or fit alongside the ecological approach. In any case, whether it is essential to enjoying or even completing a game or "merely" in the background, music is part of a video game's ecology, and thus part of shaping player interaction with the game.

References

Aarseth, E. (2004). Genre Trouble. *Electronic Book Review* 3. http://www.electronicbookreview.com/thread/firstperson/vigilant. Accessed 1 November 2012).

Calleja, G. (2011). *In-game: from immersion to incorporation*. Cambridge: MIT Press.

Clarke, E. F. (2005). *Ways of listening: an ecological approach to the perception of musical meaning*. Oxford: Oxford University Press.

Collins, K. (2008). *Game sound: an introduction to the history, theory, and practice of video game music and sound design*. Cambridge: MIT Press.

Cooke, M. (2008). *A history of film music*. New York: Cambridge University Press.

Dreyfus, H., & Kelly, S. D. (2007). Heterophenomenology: heavy-handed sleight-of-hand. *Phenomenology and the Cognitive Sciences, 6*, 45–55.

Farrow, R., & Iacovides, I. (2012). In the game? Embodied subjectivity in gaming environments. In *6th International Conference on the Philosophy of Computer Games: The Nature of Player Experience*, 29–31 January 2012, Madrid, Spain.

Gaver, W. W. (1993). How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology, 5*, 285–313.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Grimshaw, M. (2007). *The acoustic ecology of the first-person shooter*. Ph. D. The University of Waikato.

Grimshaw, M. (2008). *The acoustic ecology of the first-person shooter: the player experience of sound in the first-person shooter computer game*. Saarbrücken: VDM.

Grimshaw, M. (2012). Sound and player immersion in digital games. In T. Pinch & K. Bijsterveld (Eds.), *The Oxford handbook of sound studies*. Oxford: Oxford University Press.

Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Mahwah: L. Erlbaum.

Ihde, D. (2007). *Listening and voice: phenomenologies of sound*. Albany: State University of New York Press.

James, W. (2007). *The principles of psychology: volume 1*. New York: Cosimo.

Johnson, M. (2007). *The meaning of the body: aesthetics of human understanding*. Chicago: University of Chicago Press.

Jørgensen, K. (2008). Audio and gameplay: an analysis of PvP battlegrounds in World of Warcraft. *Game Studies, 8*.

Juul, J. (2005). *Half-real: Video games between real rules and fictional worlds*. Cambridge: MIT Press.

Kennedy, H. W. (2002). Lara Croft: feminist icon or cyberbimbo? On the limits of textual analysis. *Game Studies: International Journal of Computer Games Research*, *2*.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*, 513–549.

Krueger, J. W. (2011). Doing things with music. *Phenomenology and the Cognitive Sciences, 10*, 1–22.

Leman, M. (2008). *Embodied music cognition and mediation technology.* Cambridge: MIT Press.
Linderoth, J., & Bennerstedt, U. (2007). This is not a door: an ecological approach to computer games. In: *Situated play.* Tokyo: University of Tokyo.
Merleau-Ponty, M. (1962). *Phenomenology of perception*. New York: Humanities Press.
Miller, K. (2012). *Playing along: digital games, Youtube, and virtual performance*. Oxford: Oxford University Press.
Nattiez, J. (1990). Can one speak of narrativity in music? *Journal of the Royal Musical Association, 115*, 240–257.
Pinchbeck, D. (2009). *Story as a function of gameplay in first person shooters: an Analysis of FPS Diegetic Content 1998–2007*. Ph. D. University of Portsmouth.
Scruton, R. (1997). *The aesthetics of music*. Oxford: Clarendon.
Shultz, P. (2008). Music theory in music games. In *From Pac-Man to pop music*. Karen Collins (Ed.). Aldershot, Hampshire, [etc.]: Ashgate
Whalen, Z. (2007). Film music vs. video game music: the case of Silent Hill. In *Music, sound and multimedia*. Jamie Sexton (Ed.). Edinburgh: Edinburgh UP,
Windsor, W. L. (2004). An ecological approach to semiotics. *Journal for the Theory of Social Behaviour, 34*, 179–198.
Windsor, W. L., & de Bézenac, C. (2012). Music and affordances. *Musicae Scientiae, 16*, 102–120.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.