2008

Autopoiesis and sonic immersion: modeling sound-based player relationships as a self-organizing system.

Mark Grimshaw

University of Bolton, m.n.grimshaw@bolton.ac.uk
ABSTRACT
In previous work I have provided a conceptual framework for the design and analysis of sound in First-Person Shooter games and have suggested that the relationship between player and soundscape in such games can be modeled as an acoustic ecology. This paper develops these ideas further in the context of multiplayer First-Person Shooter games. I suggest that individual acoustic ecologies within the game combine to create a virtual acoustic ecology, of which no player is wholly aware, and that this virtual acoustic ecology may be modeled as an autopoietic (sonopoietic) system that, in part, explains and enhances player immersion in the game.

Categories and Subject Descriptors
H.5.5 [Information Interfaces and Presentation]: Sound and Music Computing – modeling, systems.

General Terms
Design, Human Factors, Theory.

Keywords
Autopoiesis, FPS games, Immersion, Self-Organizing Systems.

2. INTRODUCTION
Games such as the Doom [9], Quake [10] or Half-Life [22] series are a particular sub-genre of the FPS game known as the FPS run and gun game. These FPS games typically have a focus on violent action, in the form of killing other game characters, at the expense of long-term strategy or complex narrative. Which is not to say that strategy and narrative play no part — Half-Life 2 [22] is a case in point requiring the player to strategize in order to reach an ultimate goal over many hours of game play as opposed to Quake III Arena [10] whose purpose is simply to win each level with short bursts of frenetic, visceral activity (such levels being independent of each other in terms of objectives and strategies). Lacking a driving narrative and plot and, therefore, substantially devoid of the requirements for strategic long-term thinking, FPS games use different methods to engage the player.

With increases in computer network bandwidths and the growth of the Internet, almost all modern FPS games provide both single-player and multiplayer configurations where, in the latter case, networked players compete against, or play as a team member with, other players connected to a central game server via either a Local Area Network (LAN) or through the Internet. There are a variety of game modes variously offered by different FPS games including the free-for-all or deathmatch mode (in which players attempt to kill as many other players' characters as possible) and team-based modes such as team deathmatch and capture the flag (in which points are gained by assaulting the enemy team's stronghold, capturing their flag and transporting it safely back to the team's base).

All FPS games, as the name implies, present the player with a first-person perspective where the image presented on screen is intended to be what the player would see were he physically to take the place of his character within the game.1 The more modern FPS games have increasingly detailed and sophisticated visual and sonic environments making use of complex perspective, scaling and parallax techniques in addition to multiple audio channels, real-time mixing and acoustic shading among other features. Additionally, continuing the tradition started by Catacomb 3D [4], modern FPS games enhance this first-person perspective by the addition of a hand or pair of hands on the screen clutching a weapon that recedes perspectivally from the player into the game world. The fact that they typically respond to player input with a variety of animations (such as reloading a gun at the player's command) enhances the desired player immersion into the game world through a physically reactive identification with the game character.

One of the goals of the FPS designer is to make the player believe that he is within the game environment, that he is the character whose hands he sees before him. Back stories and promotional material always address the player in the second person singular and situate the player within the world postulated by the game: "[Y]ou take the battle to the heart of the Strogg home planet and become Earth's only hope for victory" [12]. This first stage of game immersion is supported by advances in computer technology, both hardware and software, leading to back-of-the-box claims such as "incredible graphics, and revolutionary

---

1 Albeit with some limitations such as a lack of peripheral vision. Furthermore, framed as it is within the boundaries of a visual monitor, the image on screen competes with the images from the player's environment. Some FPS games provide an optional third-person mode of play or a temporary third-person mode in particular situations.
technology combine to draw you into the most frightening and gripping first person gaming experience ever created” [11].

The FPS game affords the player the perception of moving around the game world. This illusion is achieved, both visually and aurally, by the game code creating the game's visual environment in real-time by, among other techniques, (re)positioning the game's visual objects (affording the illusion of parallax), by scaling visual objects differently (affording the illusion of visual depth) and by moving sound objects around the player as he issues movement commands. This is merely one of the ways in which the FPS game attempts to immerse the player within the virtual space of the game world but has been critiqued by Taylor:

"The very attempt to bring a player into the game space through the screen by means of a first-person point-of-view is, ironically, inconsistent because the first-person point-of-view assumes that the player herself can be caught into the structure of the game and can then be incorporated into the game space. In this way first-person perspective assumes that by enveloping the player as the player into the game space, the player becomes part of the structure of the game space [21]."

Here, the game spaces are defined by Taylor solely as visual spaces but, barring the development of holodeck-style games, there is as yet no visual enveloping of the player. That is to say, there is no suggestion that a player is physically immersed with the Cartesian co-ordinate system at the foundation of any FPS game visual spaces. I have, in other work, suggested that the FPS player is, in fact, physically immersed in a range of acoustic or resonating spaces [7] and am engaged in ongoing psychophysiological research to objectively measure levels of FPS player immersion attributable to sound [8]. Indeed the majority of sounds derive from player’s actions, either implicit (moving into different areas of the game world containing different environmental sounds) or explicit (the firing of weapons) and this derivation serves as evidence of the player’s engagement with, if not immersion in, the game.

Furthermore, in the FPS game, the player is encouraged to believe that he is immersed within an illusory 3-dimensional game world and he is, in fact, able to interact with it in significant ways. Players can see the results of their actions, navigate and explore the game world and these lead to a different game play exegesis; FPS game play is fundamentally non-linear and depends on the player for the direction it takes (within the confines of kill or be killed). The hunter and the hunted premise of these games (and discussion of an acoustic ecology must take between players that is founded) is fully encompassed in real-time by, among other techniques, (re)positioning the game's visual objects (affording the illusion of parallax), by scaling visual objects differently (affording the illusion of visual depth) and by moving sound objects around the player as he issues movement commands. This is merely one of the ways in which the FPS game attempts to immerse the player within the virtual space of the game world but has been critiqued by Taylor:

The scientific definition of ecology describes a study or an area of knowledge and acoustic ecologists define the term acoustic ecology as concerning “the relationship between soundscape and listener” [23]. In the case of the acoustic ecology of the FPS game, then, I suggest that it is the set of relationships between the player(s) and the soundscape(s) of the game. I further suggest, because the sound heard by any individual player in a multiplayer FPS game depends, to a large degree, upon the actions of other players, that the acoustic ecology of a multiplayer FPS game also includes a set of relationships between players that is founded upon sound. The importance of this sonic network between players is further stressed by reminding the reader of the hunter and the hunted premise of the FPS game. Human stereoscopic vision is worthless when confronted by the flat 2-dimensional monitor of a computer game and so stereophonic hearing becomes fundamental to the player's relationship to the game and to other players. The notion of an acoustic ecology of the FPS game (and the conceptual framework upon which it is founded) is fully described in previous work [5], [6].

Including the actions of players throughout game play encompasses the role of time in helping to both create and modify the acoustic ecology — time is a fundamental component of sound and thus any discussion of an acoustic ecology must take this into account. In the conceptual framework mentioned above, I defined the terms ideodiegetic (those sounds that any one player hears — self-produced sounds are kinediegetic, other sounds are exodiegetic) and telediegetic (those heard and responded to by a player — they are ideodiegetic for that player — but which have

3 Ecology is also often used more directly and less conceptually both in popular usage and by ecologists themselves: "We are not outside the ecology [...] we are always and inevitably a part of it” (Bateson quoted in Westerkamp [23]). This quote is given, and accepted uncritically, within the same body of writing by Westerkamp from which the previous quote is taken. Here, I use the term acoustic ecology in the second sense as a synonym for acoustic ecosystem.
consequence for another player; they are telediegetic for the second player). Teledeiegiosis, though, is a concept that also helps to explain the temporal depth of the multiplayer virtual acoustic ecology that I describe further below and can be used to explain the temporal relationships between players; telediegiegisis incorporates the notion that sounds heard in the past by one player may have the potential to later affect the game play of other players. This temporal dimension, as modeled through the concept of teledeiegisis, becomes important when I discuss player immersion and autopoiesis below.

Viewing the world of the FPS game as an ecology rather than an environment allows for a holistic view of the game of which the acoustic ecology (comprising player(s) and soundscape(s) and the relationships between them) is a part. Furthermore, it helps to explain immersion in the game through sonic means because it is a conceptual model that encompasses, and indeed is predicated upon, the player. It is a player-centred construct in which the game’s sonic environment, its soundscape, requires the presence of Böhme’s discerning player to apprehend it [1]; the trial of the soundscape, the player and the relationship between them forms the basis of the acoustic ecology.

3.2 The Virtual Acoustic Ecology

In a multiplayer FPS game, there is not just one acoustic ecology but several and, furthermore, these operate within a virtual acoustic ecology. In ecology (that is, the biological study), it is possible to study the ecology of, for example, a sandy river-bed, the ecology of a nearby kopje and the ecology of the termite mound and these may all be seen as inter-related components of the wider ecology of the Kalahari scrub. So it is with the multiplayer FPS game where each player and his soundscape form a unique entity that may be studied as an acoustic ecology by itself but where each acoustic ecology forms part of a larger whole. Like the real-world ecology, the components of such individual ecologies are not necessarily fully aware of other acoustic ecologies (the queen termite, while fulfilling her reproductive function, operates solely within her own) but they are affected by them and, indeed, may have aspects in common.

Whereas the ecologies forming the larger whole of the Kalahari ecology shift and interact at a relatively slow pace, by comparison, individual ecologies of the FPS game undergo rapid change not only in their components and position within the larger ecology but also in their interaction with other players’ acoustic ecologies. Like a snail, game characters carry their kinediegetic acoustic environments with them as they move. This strengthens the player’s identification with the character through the process of synthetic proprioception, itself aided by synesthesia, the first-person perspective of the FPS game and the player’s position as first-person auditor [7]. However, the act of moving around the game world causes the players’ acoustic ecologies to mutate as they incorporate the acoustic ecologies of other players. In the fast-paced world of the FPS game, this process is highly volatile and requires well-developed listening skills on the part of the player if his character is to survive long.

The operation and maintenance of the FPS virtual acoustic ecology may be explained through applying the principles of autopoiesis to an expanded and simplified version of the acoustic ecology model. In this model, I suggest that each player’s acoustic ecology might be viewed as the phenomenological domain of an autopoietic system (comprising the allopoietic components of player, soundscape and game engine) and that this autopoietic system itself is an allopoietic component of the autopoietic system that is the virtual acoustic ecology of the networked, multiplayer FPS game. Furthermore, autopoietic theory can provide a means to further explain a player’s immersion in his acoustic ecology.

4. AUTOPOIESIS AND IMMERSION

4.1 The Virtual Acoustic Ecology

An autopoietic system is a homeostatic organization devoted to the maintenance of that organization. External information is viewed as a perturbation to which the autopoietic system responds by compensatory processes of transformation (production and destruction of its components) to further the goal of maintaining its organization as an autopoietic system. By this definition (and focusing on a single player’s FPS game acoustic ecology), the autopoietic system comprises the FPS game system (containing hardware, the computer code and the game’s audio samples), the player and the soundscape, and the system’s phenomenological domain is the acoustic ecology. The purpose of the system is the preservation of its organization as an autopoietic system which entails the maintenance of its phenomenological domain, the game’s acoustic ecology. If that domain is defined in terms of its ability to indicate spaces, places and times in addition to indications of player activity, then the transformations of the system’s component sounds are compensations for the perturbations in the networked, multiplayer FPS game. These compensations are for the maintenance of the structures of the acoustic ecology (these structures may undergo transformations as long as they remain structures defining the acoustic ecology). If this fails, if the player is no longer able to perceive the acoustic ecology of the game, then the acoustic ecology no longer exists as a phenomenological domain and the autopoietic system has failed in its purpose and is, therefore, no longer autopoietic. Players are...

---

4 Other researchers may be interested in defining and investigating the whole game as an ecology.

5 While an extended view of FPS game play (a game level in progress) as an autopoietic system is beyond the scope of this article, it may be a task for future research. Additionally, Puterbaugh has defined the term sonopoiesis: “Sonopoietic space is the space of listening that we create through the act of listening to sound” [19]. However, this is insufficient for an understanding of the FPS game acoustic ecology as the phenomenological domain of an autopoietic system as it does not account for the player’s input and ability to trigger sounds. It is a listening space, more akin to cinema, that is less physically created or contributed to by the listener than the actively-created listening and participatory spaces found in FPS games.
no longer components of that system and therefore, I argue, no immersion takes place.

Figure 1 is a schematic of a multiplayer FPS game's virtual acoustic ecology in which five allopoietic components (comprising game engine and an individual player's acoustic ecology) are positioned. Telediegesis is used to explain the ripple of perturbations as a new player (external information) joins the game having the strongest effect upon allopoietic components within closer virtual hearing distance and a weaker effect upon those more distant. Within the virtual acoustic ecology, allopoietic components send ripples of perturbations to other allopoietic components which, when viewed as autopoietic systems, respond to this external information with compensatory transformations of their own phenomenological domains (their acoustic ecologies) thus sending forth further perturbations. The allopoietic components move throughout the space of the virtual acoustic ecology as their player components move throughout the game world, jostling with other allopoietic components and thereby deriving new exodiegetic sounds for their acoustic ecologies and contributing exodiegetic sounds to other acoustic ecologies.

For diagrammatic simplicity, only two such ripples are shown.

4.2 Immersion Through Autopoiesis

The transformations of the virtual acoustic ecology play a part in enabling the new player's immersion in the system because the process of compensation is one of inclusion. That is, the compensation takes the form of the autopoietic system's inclusion of a new allopoietic component (the player and all his kinediegetic sounds) and this inclusion immerses the player and his acoustic ecology in the virtual acoustic ecology and, as a result, immerses him in the world of the game. Furthermore, it may also be postulated that if the player, an allopoietic component of the acoustic ecology, is viewed in autopoietic terms, then the compensatory responses of the player to perturbations arising in his soundscape (giving rise to the compensatory production and destruction of sounds) also lead to immersion in the game's virtual acoustic ecology. This process of poiesis (the creation of the player's soundscape) and immersion in the game's virtual acoustic ecology is demonstrated in Figure 2 where one allopoietic component of the autopoietic system is shown.

In the diagram of Figure 2, the cyclical nature of the autopoietic system that is manifested by the player's acoustic ecology is demonstrated through a chain passing from player to game engine to soundscape back to player. The context of this acoustic ecology is as an allopoietic component of the autopoietic system that is the virtual acoustic ecology but, artificially removing it...
from this context, it may itself be assessed as an autopoietic system in its own right.

Sonic poiesis (autopoietic sonification, in other words) is the process by which the soundscape is created in response to player input both internal and external. Initially, expanding upon Böhme [1], only the player’s presence in the game is required; a willingness to engage with the system as first-person auditor [7]. In this case, it may be suggested that the player, upon becoming a willing ‘immersee’, sacrifices an autopoietic existence outside the virtual acoustic ecology to become an allopoietic component of the autopoietic ecology and this in itself is a guarantee of immersion in the circularity and homeostatic nature of the autopoietic system.

The soundscape provides feedback to the player that is filtered through a range of perceptual factors such as player experience and expectation, modes of listening and affordances [5], [6]. This in turn provokes responses from the player that are fed back to the game engine thereby continuing the cycle. The initial cycle creates the soundscape as the discerning player enters the game, which then becomes a component (with the player) of the individual acoustic ecology, and the following cycles work towards a process of equilibrium within the virtual acoustic ecology (through the process of destruction and construction — that is, the sounding and silencing of audio samples); a self-organized equilibrium that is characteristic of autopoietic systems. Furthermore, as this equilibrium is reached, the player’s immersion in the game’s virtual acoustic ecology is increased to its maximum because the ecology’s compensatory transformations involve the incorporation of external information (the player) to the point at which there is no longer any external information, only a new allopoietic component. It is likely that greater player experience (a perceptual factor influencing the meaning given to sounds in the FPS game) will decrease the number of cycles required for full immersion-equilibrium. This experience may come from a particular game or through the conventions of FPS games in general (which themselves often appropriate the expectations aroused by other forms of media such as cinema).

Perturbations in one player’s acoustic ecology are derived from the compensatory processes occurring in the game’s virtual acoustic ecology and these processes result not only from the inclusion of new players, as noted above, but also from the ongoing actions of all participating players in the game. Here, telediegesis proves useful in explaining how this occurs. Perturbations in the system may be conceived of as ripples expanding throughout the system but, as they expand, lessening in intensity (see Figure 1). This is a process that takes place over time and, to explain it as an instance of telediegesis, the following possible FPS capture the flag scenario is provided. Here, the heroine is about to save her team’s flag-carrier from the attack of an enemy soldier. At this point in time, there are at least three players in the scene. This conjunction has not happened by chance — sound has played a key role. In the case of the enemy, he has been following the flag-carrier through the twisting passageways of the game’s buildings by tracking the sound of footsteps and gunfire (navigational listening). Similarly, our heroine has been drawn to the unfolding drama by the sounds of distant battle (again, navigational listening).

The conjunction is a result of players’ responses to exodiegetic sound which, in the initial phases of the chase, would not have been heard by the opposing player. This is telediegesis; the reaction by the enemy to the sounds of the flag-carrier having later consequence for the heroine (and her team) such that she comes face-to-face with the enemy and is therefore able to save her flag-carrying team-member by killing that enemy. Thus, telediegesis is a perturbation rippling through the virtual acoustic ecology from the enemy’s acoustic ecology to the player’s and vice-versa. Both these autopoietic systems compensate for these telediegetic perturbations through the manifestation of new, common sounds (the production of new components) thereby contributing to the maintenance of the phenomenological domains that are the players’ acoustic ecologies and the game’s virtual acoustic ecology. The destruction of components is exemplified by the removal of the killed enemy’s sounds from the heroine’s acoustic ecology.

4. CONCLUSION

The suggestion that the game’s virtual acoustic ecology is the phenomenological domain of an autopoietic system needs to be treated with care. As Maturana and Varela state, “autopoiesis generates a phenomenological domain, this is cognition” [16]. Thus, the existence of the virtual acoustic ecology presupposes cognition on the part of the autopoietic system where cognition includes the ability to acquire, store, retrieve and use knowledge. The FPS game engine (and associated game components such as hardware and computer memory, for example), in its capacity to acquire and store game information, to retrieve and use that information for the maintenance of its organization and ‘phenomenological domain’, is therefore, by this definition, cognate. The logical autopoietic conclusion then, is that the FPS game in progress is living because it is autopoietic and autopoiesis is the sole requirement for life (according to autopoietic theory, “autopoiesis is necessary and sufficient to characterize the organization of living systems” [16]). It is not the purpose of this article to argue that the FPS game’s acoustic ecology is a living being nor does it. But it is an interesting topic for future debate because here it has been demonstrated that there are many aspects of the virtual acoustic ecology that may be described and explained through autopoietic theory such that it becomes possible to tentatively suggest that the emergent system [13] of the virtual acoustic ecology is an autopoietic system.

There is a stronger argument to make that each individual player’s acoustic ecology is a phenomenological domain arising out of a cognate autopoietic system because such a system includes the sentient player. Furthermore, the player is potentially fully aware

That is, in ecological terms, the climax of the ecology.

8 Particularly as the game engine is not able to sense sound or, indeed, any of the acoustic phenomenological domain to which it contributes.
of all elements of his acoustic ecology and thus it may justly be described as his (acoustic) phenomenological domain. However, in the virtual acoustic ecology, though it contains players, and though each player's acoustic ecology may share common sounds with others, each player is unaware of the totality of another's acoustic ecology and, furthermore, unaware of the totality of the virtual acoustic ecology. Swinging the argument the other way, though, this article has argued that the majority of sounds in the FPS game arise out of, and are evidence therefore, of player actions and, thus, player presence in the game. The virtual acoustic ecology can be said to be maintained almost entirely by the active participation of players, the willing ‘immersees’ previously described, reacting to other allopoietic components of the system. The virtual acoustic ecology can perhaps be seen as a phenomenological domain produced by multiple, sentient players each of whom perceive just a portion of it. However, as with Jean-Luc Picard’s ineluctable pondering on the Borg Collective, a discussion of the FPS game (acoustic ecology) as life-form is a subject for future debate.

5. REFERENCES