TOM LANGHORST
Fontys University of Applied Science
t.langhorst@fontys.nl

The Sound of a Serve Toss

An Informational View on the Gameworld Interface as Sonic Interface Design

ABSTRACT

In this article, Jørgensen's three-layer model of the gameworld interface is used to analyze the practice of game sound design and to describe the informational role of game audio design, from the perspective of *Sonic Interface Design* (SID). First the relationship between visual and audio realism is discussed. Then three layers in Jørgensen's model (controllers, WIMPs elements, and gameworlds) are explored in relation to the semiotic analytical model of C. S. Peirce. By addressing the sound of the serve toss in games such as *Pong* (1972) and *Wii Sports Tennis* (2006), the article investigates how sound design can support realism, hypermediacy, and immediacy in varying contexts. Approaching game sound design from an information-based concept (the SID) as well as a holistic model (as gameworld interface), the aim here is to improve an understanding of the creative potential and extent of game audio design.

KEYWORDS: *Game Sound Design, Sonic Interface Design, Informational Spaces, Music Cognition, Semiotics, Persuasive Design*

INTRODUCTION

Kristine Jørgensen (2013a, 2013b) provides a fresh view on gameworlds by conceptualizing games as informational spaces working through three layers of an interface: the game controller, the WIMP elements (windows, icons, menus, and pointers) and the gameworld. Here this model is used to consider game sound design defined as *Sonic Interface Design* (SID) which "explores ways in which sound can be used to convey information, meaning, aesthetic and emotional qualities in interactive contexts." (Serafin et al. 2011, p.87). *As Collins notes* (2008), in the game *Pac-Man* (1980, Namco, Japan), "(t)ypically the music only played when there was no game action, since any action required all of the system's available memory" (p. 9). Considering aspects of immersion, it may seem like a disadvan-

tage when, in early video games, it was not possible to combine sound effects and music because of limitations in available memory space; however, the omission helped players to understand the mechanics of the game better. The examples of *Pac-Man*'s "bite" and *Pong*'s serve toss sound can be considered to understand how different audiences may have different expectations of what a "realistic", "believable", or "appropriate" sound effect may be in a game.

When we look at one of the core mechanics of *Pac-Man* ("eat blocks"), neither the sound of taking bites nor the animation can be regarded as realistic. The concept of the mechanic is suggested by the up and down movement of *Pac-Man*'s mouth and a sound effect that vaguely resembles the sound of a bite. The perception of the sound effect as "bite" is, like all game audio (Jørgensen 2008), highly contextual. Furthermore, following the semiotic approach of C.S. Peirce (see, for example, Zalta, 2010)[1], a sound effect may be regarded as an iconic sign, a representation that resembles qualitative features of the object. Peirce introduced ten types of signs of which three types of object-sign relations, commonly called "modes", are applied to this discussion: iconic, indexical and symbolic. When the sign has an analog physical connection between the signifier and its object, then the sign is an indexical sign. And if the sign utilizes arbitrary convention, habit, or social rule or law according to which the signifier indicates the object, it is a symbolic sign Even though there is some allusion to a ping pong ball, the sound effects of *Pong* (1972) is even less realistic than *Pac-Man*'s. Even more, the visual design of *Pong*'s ball, a square set of low-resolution pixels, undermines the concept "ball" in the iconic mode (one expects a real ball to be round), but as a symbolic sign, it works: players do not seem to mind a lack of realism, as long as it functions as a type of ball in the game. As the gameplay of the ball-game is preserved, the minimalism of both sonic and visual representations balance each other without disturbing the understanding of the game mechanics.

Awareness of aspects of sonic design in relation to the gameworld and players' expectations can help make the production process of game sound design methodical and effective, but it also implies a more holistic approach to game sound design than what the one that dominates current game studies and game development.

Modes of representation differ according to the demands of context and game. If we were to use Pong's sound effects in the sound design of the much more realistic *Wii Sports Tennis* (2006), the balance would be disturbed. Each year when I introduce my students to game sound design and show them a gameplay movie of *Wii Sports* with the sound effects of *Pong* they are amused by the disturbance of their expectations of a more realistic sound (Huron 2006). Indeed, the relatively more sophisticated realism of *Wii Sports* shapes different expectations in the player than *Pong*'s sound effects.

There is, however, one exceptional sound in *Wii Sports Tennis* that outbalances its visual and sound realism: the serve toss.

This sound seems to have escaped from a cartoon soundtrack and sneaked its way into the *Wii Sports Tennis* game. The question is: why? The movement

1. Within Peirce's analytical approach, a sign consists of three inter-related elements: a sign (signifier), an object (signified), and an interpretant who makes sense of a sign. In the context of sound design, we can say that the sound itself is the sign element (signifier), whereas the object is what the sound stands for; so, in the Pac-Man example, the object can be described as "bite". The interpretant is best thought of as the understanding that the listener has of the sign–object relation. The process of game sound design can be related to these three elements when we consider its resulting product, the sound, as the sign which is based on the designer's applied theoretical and practical knowledge of sound design in terms of the sign's object and user test-based validation of the sound as the interpretant. Therefore, developing game sound design knowledge implies a focus on describing design patterns that can be deducted from knowledge of the relationship between the signifier and its object.

of a ball in the air (serve toss) can theoretically be considered but not practically be perceived as a sound for psychoacoustical reasons. Therefore, the serve toss sound can be regarded as a symbolic sign based solely on conventions. The upwards glissando (raise in pitch over time) represents the upward movement of the ball. Here, the sound of the serve toss is a good example of sonification, "the technique of rendering sound in response to data and interactions.' (Hermann et al. 2011, p.1). Sonification of movement in a game mechanic is not limited to this example alone.

The jump of Mario in *Super Mario Bros.* (1985) is another illustration of the same kind of sonification.

Such aspects impact on the gameworld as an informational space. Again, the player's interaction with a game can be described as a three-layer interface consisting of the game controller, WIMP (Windows-Icons-Menus-Pointers) and gameworld (Jørgensen 2013a, 2013b). According to Jørgensen, a gameworld as interface can be designed between the two opposing approaches of superimposing and integrating media, representing either a hypermediacy style of visual representation, the goal of which is "to remind the viewer of the medium", or an immediacy style of visual representation, which aims to "make the viewer forget the presence of the medium" (Bolter & Grusin 1998: pp. 272-273). Implementing the serve toss as a more or less "realistic" sound, while considering the affordances of context, is part of this strategy. Jørgensen explicitly includes audio as an important element of the gameworld as interface. The following part of this paper discusses how sound design can support these layers and design approaches in varying semiotic modes.

THE WIMP LAYER

The WIMP layer can be situated in the game's (graphical) user interface, or (2D) Head-Up Display, typically using windows, menus and the gameworld when (3D) pointers and icons are added inside the world. The main purpose of the WIMP layer is to provide the player with game system information (Jørgensen 2013b). In the WIMP layer, sound effects are most commonly used in the iconic mode when they support click actions related to the WIMP elements. In this case, the sound effect resembles the real-world sounds of our interaction with, for example, switches, buttons, and paper turns.

The click action in the WIMP layer can also be supported by sound effects that "look ahead" or "summarize" the (next) game state by representing the meaning of the state as a symbolic sign. In the app version of the game Risk, the WIMP layer sound effects contain elements of "arming riffles", "steps of army boots", and the "soundscape of a battlefield" to represent (part of) the meaning of the game states.

When the WIMP layer is used to indicate the outcome of our gameplay, for example when we score in Pong, the sound effects can be used to evaluate the gameplay. In that case, the design principle or design pattern of the evaluated

sound effect is important to establish its semiotic mode. Comparing the evaluative audio signs (success and failure) of *Pong* and *Wii Sports Tennis* we can say that *Pong's* sound effects are symbolic, based on conventions, while the sound effects of *Wii Sports Tennis* are iconic since they resemble the sound of a crowd[2]. Nevertheless, such effects can also be considered as indexical signs, signs (in this case the sound recordings) that result from, are analogous to, something that happens in the real world This explains why such sound effects are so successful in communicating the meaning of success and failure (Langhorst 2014).

THE GAMEWORLD LAYER

In the gameworld layer of Jørgensen's model, a sound effect may exist in all three semiotic modes, iconic, symbolic and indexical, and can shift from one mode to another, and appear in combination. The *Wii Sports Tennis* serve toss sound is symbolic as a sign representing the event of a ball thrown in the air[3]. It is, however an index sign in relation to the data it represents (the change in the height of the ball). This is crucial for the concept of sonification where the sound must represent data in a one to one relationship to provide the player with accurate information.

Jørgensen (2013b) notices a semiotic mode shift in the gameworld that is caused by learning: "(w)hen we have played the game enough to learn how it works, however, signs that were formerly mysterious change character and become familiar and recognizable. Now the signs become representative of specific events in the gameworld" (p. 79). When it comes to sounds that hold a strong first/third person relationship between game and player, such as footsteps, the idea of internalized signs shifting from the iconic mode into the indexical mode is substantiated by arguments from recent neuropsychological research that "consistently suggested that the brain processes the sounds of actions made by an agent differently from other sounds" (Serafin et al. 2011, p. 89).

The indexical mode-shift of audio-signs can be compared with the acceptance of the (unedited) photographic image as an indexical sign (Barthes 1977) and explained as the perception of realism. An improved understanding of perception, combined with improved technology to create virtual worlds, makes it possible to convert the unreal, or the virtual, into perceived reality, *or modality*. The latter concept "refers to the degree of truth assigned to a given sound event" (van Leeuwen, 1999, p. 180). In short, the codes and conventions of game audio, sound effects, and game music, interact with other dimensions of the game to produce a believable, immersive experience for the player.

AUDIO EVENT DISTRIBUTION

The function of sound events in games is to provide the player with important information to understand the game and its game mechanics through detailed sound design; sounds can do so in either an integrated or superimposed manner. Sonic interface design provides the possibility to create highly informative game-

2. The evaluating sound effects of *Wii Sports Tennis* have become part of the gameworld layer but nevertheless provide the player with game system information. Game sound design can help to integrate game system information into the gameworld layer.

3. The relationship between the upward moving sound and ball is much more complex than one might think at first. Since pitch is perceived logarithmically and is therefore described as pitchΔtime = log2 freqΔtime , the linear frequency rise from 200Hz to 600Hz of the SFX will be perceived as a logarithmic pitch change from approx. G to d". The position of the ball is a parabolic function that can be describes as positionΔtime = (V0$\star\Delta$T) + (1/2A$\star$ ΔT2) , where ΔT is delta time, A is the world's gravity (- 9.8m/s2) assuming that the Wii sports world behaves like ours, and V0 is the initial speed of the ball. As the ball in Wii Sports tennis rises until its velocity becomes 0 and the ball reaches its highest position (the optimum to hit it!), the curve of the logarithmic pitch function and the parabolic ball position function are very much alike. This is due to the characteristic of the two functions but also due to the variable values (glissando start and ending pitches, and SFX length) make the two curves more alike under the condition that the player needs to hit the ball while it is rising. Thus, however not mathematically identical, in effect the two curves may be perceived as equal, and therefore the server toss SFX is perceived as the sonification of the position of the ball.

worlds, revealing game system information in an integrated manner, whereas dominance of visual information tends to lead to additional superimposed media.

Gameworlds with integrated audio events are alike our real-life environment and listening to this informational space can, therefore, be described as "(e)very-day listening [, which] is the experience of hearing events in the world rather than sounds per se" (Gaver 1993, p. 285). Gaver's framework explains how we perceive nonmusical sounds as a distribution of events and describe them not by their physical characteristics but by their source; for example, we describe a sound as "the siren of a passing police car" rather than that we summarize the physical specifications of the Doppler effect. Furthermore, Gaver shows how tolerant we are regarding the physical accuracy of a sound we hear: we have no problem recognizing footsteps regardless whether they sound on the street or in a large church despite that this significantly changes the physical characteristics of the sound. Game sound design can benefit from this tolerance to shift sound from accuracy to immersive experience. Summers' (2016) analysis of five different game types within the race game genre shows they each have their own specific relationship with how immersion serves realism, not only by facing the challenge of combining music and the dynamical sound of a racecar but also by the need to balance sound design in regard of realism, genre, and immersion.

Not only everyday sounds but also music can be perceived as "events" (Buxton et al. 1994). While listening to music involves a significant amount of cognitive processing within a complex auditory system, prediction seems to be an important aspect of this process (Huron 2006, Zatorre and Salimpoora 2013). Game music has been a topic of research quite some time, which seems to imply that interacting with game music is different than listening to music in general (Collins 2013). For example, *dynamic audio* "reacts both to changes in the game-play environment, and/or to actions taken by the player" (Collins 2008, p. 4), which is essential for game audio and for the immersive function of game music (Phillips, 2014), In relation to game music, this is often referred to as *adaptive music*. In a game with adaptive music, the game system's logic needs to evaluate the gameplay and accordingly change the game state resulting in a change in the music. This process implies two different levels of cognitive processing for the player, initiated by unpredictable event-based decisions of the game system's logic. This causes an additional demand on the cognitive process during listening to music. Firstly, in terms of perception––at the first level the musical composition will change according to the characteristics of the new game state. Typically, this involves changes in the perceivable emotions and/or referential changes in the music such as the use of themes, *leitmotiv*, and the *idée fixe*, to support the game's narrative (Phillips, 2014). Secondly, concerning decoding––at the second level the player needs to decode the "argument" that initiated the game state change to perform in line with the adjusted level challenges and goals. In an ideal informational space, the design characteristics and design patterns of the music matches the argument of the game system's logic state change. Thus, for

example, a sudden attack to the playable character by in-game enemies results in a matching rise in arousal characteristics of the music, and the identification of a specific enemy involves the introduction of a theme or leitmotiv.

> To understand the full potential of sound-driven game control we need to look closer at the ways in which music can evoke motion through rhythm and expression.

THE SERVE TOSS RELOADED: THE CONTROLLER LAYER

The use of the serve toss sound in *Wii Sports Tennis* may be related to the absence of an actual ball (which one can physically strike), as an alternative cue is needed to initiate the player's action. Therefore, the purpose of the sound of the serve toss is to trigger the player's movement. According to Fogg's (2016) behavioral model, in addition to motivation and ability, triggers are the key element in persuasive design; "without a trigger, the target behavior will not happen" (online). Furthermore, triggers are especially interesting for game (sound) designers since they are the only element in Fogg's model that can be designed. The example of *Wii Sports Tennis* demonstrates that, at the game controller layer of Jørgensen's model, sound can be used successfully for changes in the player's behavior, resulting in a movement. This kind of kinetic interaction related to sound is well known from the music game or rhythm-action game genre that includes games such as *Guitar Hero* (2005) or *Rock Band* (2007).

"Rhythm-action games are video games in which the player must respond in some way to the rhythm or melody being presented, either through repeating the same melody or rhythm by pressing buttons (with hands or feet), or kinetically responding in some way to the rhythm, often using specially designed controllers" (Collins 2008, p.).

In general kinetic and sensory interaction, the haptic modality is regarded as an important aspect of interaction with virtual worlds (Mihelj, Novak & Begus, 2013). However, "the haptic modality is currently underutilized and poorly understood as a design material in game design" (Nordvall 2014, p.1). Nevertheless, sound as a trigger for behavior, as well as the rhythm-action game genre, shows the potential of game sound design at the controller layer. Therefore, to utilize sound as a trigger for motoric actions or expression seems like a valuable opportunity for new sound-driven game mechanics and controllers that go beyond the tradition of the music game genre.

To understand the full potential of sound-driven game control we need to look closer at the ways in which music can evoke motion through rhythm and expression. Firstly, concerning rhythm, movement and musical rhythm are closely connected, share neurological pathways and "indeed the close connection between music and dance suggest that musical rhythm might have evolved from rhythmic movement" (Trainor & Zatorre 2009, p.178). Some game mechanics are based on the rhythmical input of the player, as it happens in *Patapon* (2007) and in *Beat Sneak Bandit* (2012).

Secondly, in regard to expression, this also contains motion as one of its sub-components; according to Juslin & Timmers, 2010), also musical performance expression can be "conceived of as a multi-dimensional phenomenon that can be decomposed into subcomponents that make distinct contributions to the aesthetic impact of a performance "(p.454),. Unfortunately the use of expressive musical motion hardly has been explored outside the music game genre but still shows great potential if we look at the *Smart String* instrument of *Apple's Garageband app.*

Even though the physical interaction is with a limited haptic glass screen, players nevertheless intuitively adjusts their movements; they touch and impact with the screen in order to generate the expressive motion matching *pizzicato*, *marcato* and (speed dependent) dynamic sustained string sounds.

CONCLUSIONS

Game sound designers strive to design sounds that can inform the player, trigger behavior or evoke emotions. Therefore, the success of game sound design benefits from the use of design patterns. Studies in semiotics, cognitive neuroscience, psychoacoustics, affective prosody, and music psychology offer precious perspectives to understand these processes. Not only does the layered interface model of Kristine Jørgensen help game sound designers in choosing the location of their intervention with precision, it also helps them to optimize the balance between expectations of realism, hypermediacy, and immediacy of the style needed for any given game. Sonic design proves to be effective in integrated design, providing the player with information integrated into the gameworld. In the process of sound design, knowledge of design patterns is an important element for the development and production of sound. Awareness of aspects of sonic design in relation to the gameworld and players' expectations can help make the production process of game sound design methodical and effective, but it also implies a more holistic approach to game sound design than what the one that dominates current game studies and game development. Not only does a broader, and integrated, sonic approach to game design help to better understand the full potential and extent of game audio, but it also offers further creative opportunity in game design for the future.

REFERENCES

Barthes, R. (1977). *Image-Music-Text.* London: Fontana.

Boltor, J. and Grusin, R. (1998). *Remediation, Understanding New Media.* Cambridge (MA), London: The MIT Press.

Buxton, W. Gaver, W. and Bly, S. (1994). *Everyday listening. In: Buxton*, W. Gaver, W. and Bly, S. The use of Non-Speech Audio at the Interface [online] Available at http://www.billbuxton.com/Audio.TOC.html Chapter 5

Collins, K. (2008). *Game Sound – An introduction to the History, Theory and Practice of Video Game Music and Sound Design*. Cambridge MA & London: The MIT Press.

Collins, K. (2013). *Playing with Sound – A Theory of Interacting with Sound and Music in Video Games*. Cambridge (MA), London:The MIT Press.

Fogg, BJ. (2016). *The behavior model. [online]* Available at: http://www.behaviormodel.org/index.html

Gaver, W. (1993). *How do we hear in the World. Exploration in Ecological Acoustics. In: Ecological Psychology. 5(4), 285-313.* Lawrence Erlbaum Associates, Inc.

Hermann, T. Hunt, A. Neuhoff, J. (2011). *Introduction. In: Hermann, T. Hunt, A. Neuhoff*, J. ed. *The Sonification Handbook. Berlin: Logos Verlag Berlin GmbH. pp.1-7.*

Huron, D. (2006). *Sweet Expectations: Music and the Physiology of Expectation*. Cambridge (MA), London: The MIT Press.

Jørgensen, K. (2008). *Audio and Gameplay: An Analysis of PvP Battlegrounds in World of Warcraft. [online]* Available at:http://gamestudies.org/0802/articles/jorgensen Game Studies Journal. Vol. 8 issue 2 December 2008.

Jørgensen, K. (2013a). *Lecture at the Philosophy of Computer Games Conference*, Bergen 2013. [online] Available at: https://www.youtube.com/watch?v=_i_–kliuRPg

Jørgensen, K. (2013b). *Gameworld Interfaces*. Cambridge MA & London: The MIT Press.

Juslin, P.N. & Timmers, R. (2010) . *Expression and Communication of Emotion in Music Performance*. In: Juslin, P.N. Sloboda, J.A. eds. *The Oxford Handbook of Music and Emotion*. London, New-York: Oxford University Press. pp. 453–489.

Langhorst, T. (2014). *The Unanswered Question of Musical Meaning – a cross-domain approach*. In: . Collins, K. Kapralos, B. & Tessler, H. eds. *The Oxford Handbook of Interactive Audio*. London, New-York: Oxford University Press.. pp. 95-116.

Mihelj, M., Novak, D. & Begus, S. (2013). Haptic Modality in Virtual Reality. In: Mihelj, M. Novak, D. Begus, S. eds. *Virtual Reality Technology and Applications*. (pp 161-194). Dordrecht, the Netherlands: Springer Science+Business Media.

Nordvall, M. (2014). *The Sightlence Game: Designing a Haptic Computer Game Interface*. In: Proceedings of the 2013 DiGRA International Conference: DeFragging Game Studies. [online] Available at http://www.digra.org/wp-content/uploads/digital-library/paper_473.pdf.

Phillips, W. (2014). *A Composer's Guide to Game Music*. Cambridge MA & London: The MIT Press.

Serafin, S. Franinovic, K. Hermann, T. Lemaitre, G. Rinott, M. Rocchesso, D. (2011). *Sonic Interaction Design*. In: Hermann, T. Hunt, A. Neuhoff, J. eds. The Sonification Handbook. Berlin: Logos Verlag Berlin GmbH. Chapter 5.

Summers, T. (2016). Understanding Video Game Music. Cambridge: Cambridge University Press.

Trainor, L. and Zatorre, R. (2009). The Neurobiological Basis of Musical Expectations. In: Hallam, S. Cross, I. and Thaut, M. eds. Handbook of Music Psychology. (pp 171-183). London, New-York: Oxford University Press.

Van Leeuwen, T. (1999). Speech, Music, Sound. Basingstoke: Palgrave Macmillan.

Zalta, E. (2010). Peirce's Theory of Signs. Stanford CA: Stanford University. [online] Available at https://plato.stanford.edu/entries/peirce-semiotics/

Zatorre, R. Salimpoora, V. (2013). From perception to pleasure: Music and its neural substrates. PNAS Direct Submission. [Online] Available at http://www.pnas.org/content/110/Supplement_2/10430.abstract.

LUDOGRAPHY

Beat Sneak Bandit. (2012). Simogo

Garage band. Apple Corp.

Guitar Hero.(2005) RedOctane, Activision

Pac-Man. (1980). Namca.

Patapon. (2007). Pyramid SCE Japan Studio.

Pong. (1972). Atari.

Rock Band. (2007) MTV Games, Electric Arts (AE)

Risk. Electronic Arts (EA)

Super Mario Bros. (1985). Nintendo.

Wii Sports. (2006). Nintendo.

Wii Sports Tennis. (2006). Nintendo.

AUTHOR'S INFO:

Tom Langhorst is a lecturer and researcher at Fontys University of Applied Science – school of ICT, Eindhoven, the Netherlands. His work and research focus on the crossover between (persuasive) design and technology. Tom is educated as musician, music theorist and composer and previously worked in the video games, entertainment and advertisement industry, as well as interaction designer for product design and innovation.