6 | The Triple Lock of Synchronization

K. J. DONNELLY

Contemporary audiovisual objects unify sound and moving image in our heads via the screen and speakers/headphones. The synchronization of these two channels remains one of the defining aspects of contemporary culture. Video games follow their own particular form of synchronization, where not only sound and image, but also player input form a close unity.[1] This synchronization unifies the illusion of movement in time and space, and cements it to the crucial interactive dimension of gaming. In most cases, the game software's 'music engine' assembles the whole, fastening sound to the rest of the game, allowing skilled players to synchronize themselves and become 'in tune' with the game's merged audio and video. This constitutes the critical 'triple lock' of player input with audio and video that defines much gameplay in digital games.

This chapter will discuss the way that video games are premised upon a crucial link-up between image, sound and player, engaging with a succession of different games as examples to illustrate differences in relations of sound, image and player psychology. There has been surprisingly little interest in synchronization, not only in video games but also in other audiovisual culture.[2] In many video games, it is imperative that precise synchronization is achieved or else the unity of the gameworld and the player's interaction with it will be degraded and the illusion of immersion and the effectiveness of the game dissipated. Synchronization can be precise and momentary, geared around a so-called 'synch point'; or it might be less precise and more continuous but evincing matched

---

[1] With some games, a slight lack of synchrony between sound and image, between a player's input and the illusion of the event on screen, can be tolerable. For instance, with point-and-click narratives or detection games the player can mentally compensate for the discrepancy. However, the overwhelming majority of video games require rapid response to player input. What is sometimes called 'input lag' or 'latency', when a button is pressed and the in-game response to that activity is not immediate, can utterly ruin the gaming experience.

[2] I have discussed this in detail in relation to film and television in my book *Occult Aesthetics: Synchronization in Sound Film* (New York: Oxford University Press, 2014). Other relevant writing includes Jeff Rona, *Synchronization: From Reel to Reel: A Complete Guide for the Synchronization of Audio, Film and Video* (Milwaukee, WI: Hal Leonard Corporation, 1990) and Michael Sweet, *Writing Interactive Music for Video Games: A Composer's Guide* (London: Addison Wesley, 2014), 28–9.

dynamics between music and image actions; or the connections can be altogether less clear. Four types of synchronization in video games exist. The first division, precise synchronization, appears most evidently in interactive sounds where the game player delivers some sort of input that immediately has an effect on audiovisual output in the game. Clearest where diegetic sounds emanate directly from player activity, it also occurs in musical accompaniment that develops constantly in parallel to the image activity and mood. The second division, plesiochrony, involves the use of ambient sound or music which fits vaguely with the action, making a 'whole' of sound and image, and thus a unified and immersive environment as an important part of gameplay. The third strain would be music-led asynchrony, where the music dominates and sets time for the player. Finally, in parallel-path asynchrony, music accompanies action but evinces no direct weaving of its material with the on-screen activity or other sounds.

Synching It All Up

It is important to note that synchronization is both the technological fact of the gaming hardware pulling together sound, image and gamer, and simultaneously a critically important psychological process for the gamer. This is central to immersion, merging sensory stimuli and completing a sense of surrounding ambience that takes in coherently matched sound and image. Now, this may clearly be evident in the synchronization of sound effects with action, matching the world depicted on screen as well as the game player's activities. For instance, if we see a soldier fire a gun on screen we expect to hear the crack of the gunshot, and if the player (or the player's avatar) fires a gun in the game, we expect to hear a gunshot at the precise moment the action takes place. Sound effects may appear more directly synched than music in the majority of cases, yet accompanying music can also be an integrated part of such events, also matching and directing action, both emotionally and aesthetically. Synchronization holds together a unity of audio and visual, and their combination is added to player input. This is absolutely crucial to the process of immersion through holding together the illusion of sound and vision unity, as well as the player's connection with that amalgamation.

   Sound provides a more concrete dimension of space for video games than image, serving a crucial function in expanding the surface of its flat images. The keystones of this illusion are synch points, which provide a structural relationship between sound, image and player input. Synch points unify the

game experience as a perceptual unity and aesthetic encounter. Writing primarily about film but with relevance to all audiovisual culture, Michel Chion coined the term 'synchresis' to describe the spontaneous appearance of synchronized connection between sound and image.[3] This is a perceptual lock that magnetically draws together sound and image, as we expect the two to be attached. The illusory and immersive effect in gameplay is particularly strong when sound and image are perceived as a unity. While we, the audience, assume a strong bond between sounds and images occupying the same or similar space, the keystones of this process are moments of precise synchronization between sound and image events.[4] This illusion of sonic and visual unity is the heart of audiovisual culture. Being perceived as an utter unity disavows the basis in artifice and cements a sense of audiovisual culture as on some level being a 'reality'.[5]

The Gestalt psychology principle of isomorphism suggests that we understand objects, including cultural objects, as having a particular character, as a consequence of their structural features.[6] Certain structural features elicit an experience of expressive qualities, and these features recur across objects in different combinations. This notion of 'shared essential structure' accounts for the common pairing of certain things: small, fast-moving objects with high-pitched sounds; slow-moving music with static or slow-moving camerawork with nothing moving quickly within the frame, and so on. Isomorphism within Gestalt psychology emphasizes a sense of cohesion and unity of elements into a distinct whole, which in video games is premised upon a sense of synchronization, or at least 'fitting' together unremarkably; matching, if not obviously, then perhaps on some deeper level of unity. According to Rudolf Arnheim, such a 'structural kinship' works essentially on a psychological level,[7] as an

[3] Michel Chion, *Audio-Vision: Sound on Screen*, ed. and trans. Claudia Gorbman (New York: Columbia University Press, 1994), 5.

[4] Chion's synchresis matches the ideas of Lipscomb and Kendall, both of which note perceptual 'marking' by synch points. S. D. Lipscomb and R. A. Kendall, 'Sources of Accent in Musical Sound and Visual Motion', in the *Proceedings of the 4th International Conference for Music Perception and Cognition* (Liege: ICMPC, 1994), 451–2.

[5] Frans Mäyrä notes the surface and coding reality beneath. He discusses the 'dual structure of video games', where 'players access both a "shell" (representational layers) and the "core" (the gameplay)'. Frans Mäyrä, 'Getting into the Game: Doing Multidisciplinary Game Studies', in *The Video Game Theory Reader 2*, ed. Bernard Perron and Mark J. P. Wolf (New York: Routledge, 2008), 313–30 at 317.

[6] Rudolf Arnheim, 'The Gestalt Theory of Expression', in *Documents of Gestalt Psychology*, ed. Mary Henle (Los Angeles: University of California Press, 1961), 301–23 at 308.

[7] Rudolf Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye* (Los Angeles: University of California Press, 1974), 450.

indispensable part of perceiving expressive similarity across forms, as we encounter similar features in different contexts and formulations. While this is, of course, bolstered by convention, it appears to have a basis in primary human perception.[8]

One might make an argument that many video games are based on a form of spatial exploration, concatenating the illusory visual screen space with that of stereo sound, and engaging a constant dynamic of both audio and visual movement and stasis. This works through isomorphism and dynamic relationships between sound and image that can remain in a broad synchronization, although not matching each other pleonastically blow for blow.[9] A good example here would be the first-person shooter *Quake* (1996), where the player sees their avatar's gun in the centre of the screen and has to move and shoot grotesque cyborg and organic enemies. Trent Reznor and Nine Inch Nails' incidental soundtrack consists of austere electronic music, dominated by ambient drones and treated electronic sounds. It is remarkable in itself but is also matched well to the visual aspects of the gameworld. The player moves in 3-D through a dark and grim setting that mixes an antiquated castle with futuristic high-tech architecture, corridors and underwater shafts and channels. This sound and image environment is an amalgam, often of angular, dark-coloured surfaces and low-pitched notes that sustain and are filtered to add and subtract overtones. It is not simply that sound and image fit together well, but that the tone of both is in accord on a deep level. Broad synchronization consists not simply of a mimetic copy of the world outside the game (where we hear the gunshot when we fire our avatar's gun) but also of a general cohesion of sound and image worlds which is derived from perceptual and cognitive horizons as well as cultural traditions. In other words, the cohesion of the sound with the image in the vast majority of cases is due to structural and tonal similarities between what are perhaps too often approached as utterly separate channels. However, apart from these deep- (rather than surface-) level similarities, coherence of sound and image can also vary due to the degree and mode of synchronization between the two.

[8] For further reading about games and Gestalt theory consult K. J. Donnelly, 'Lawn of the Dead: The Indifference of Musical Destiny in Plants vs. Zombies', in *Music In Video Games: Studying Play*, ed. K. J. Donnelly, William Gibbons and Neil Lerner (New York: Routledge, 2014), 151–65 at 160; Ingolf Ståhl, *Operational Gaming: An International Approach* (Oxford: Pergamon, 2013), 245; and Mark J. P. Wolf, 'Design', in *The Video Game Theory Reader 2*, ed. Bernard Perron and Mark J. P. Wolf (London: Routledge, 2008), 343–4.

[9] In other words, rather than 'Mickey-Mousing' (the music redundantly repeating the dynamics of the image activity), there is only a general sense of the music 'fitting' the action.

Finger on the Trigger

While synchronization may be an aesthetic strategy or foundation, a principle that produces a particular psychological engagement, it is also essentially a technological process. Broadly speaking, a computer CPU (central processing unit) has its own internal clocks that synchronizes and controls all its operations. There is also a system clock which controls things for the whole system (outside of the CPU). These clocks also need to be in synchronization.[10] While matters initially rest on CPU and console/computer architecture – the hardware – they also depend crucially on software.[11]

The principle of the synch point where image events and player inputs trigger developments in the music is a characteristic of video game music. Jesper Kaae discusses video games as 'hypertext', consisting of nodes and links, which are traversed in a non-linear fashion.[12] 'Nodes' might be understood as the synch points that underpin the structure of interactive video games, and have particular relevance for the triple lock of sound and image to player input. Specific to video games is how the disparate musical elements are triggered by gameplay and combined into a continuum of coherent development for the player's experience over time. Indeed, triggering is the key for development in such a non-linear environment, set in motion by the coalescence of the player's input with sound and image elements. A player moving the avatar into a new room, for example, can trigger a new piece of music or the addition of some musical aspects to existing looped music.[13] Michael Sweet notes that triggered musical changes can alter emotional state or general atmosphere, change the intensity of a battle, indicate a fall in the player's health rating, indicate an enemy's proximity and indicate successful completion of a task or battle.[14]

Triggered audio can be a simple process, where some games simply activate a loop of repeated music that continues ad infinitum as accompaniment to the player's screen actions, as in *Tetris* (Nintendo, 1989). However,

---

[10] See Randall Hyde, *The Art of Assembly* (N.p.: Randall Hyde, 1996), 92–6.

[11] MMORPGs (Massively Multiplayer Online Role-Playing Games) and games for multiple players require effective synchronization to hold the shared gameworld together.

[12] Jesper Kaae, 'Theoretical Approaches to Composing Dynamic Music for Video Games', in *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*, ed. Karen Collins (Aldershot: Ashgate, 2008), 75–92 at 77.

[13] Richard Stevens and Dave Raybould, *The Game Audio Tutorial: A Practical Guide to Sound and Music for Interactive Games* (London: Focal Press, 2011), 112.

[14] Sweet, *Writing Interactive Music*, 28.

triggered audio often sets in train more complex programs where music is varied.[15] Karen Collins effectively differentiates between types of triggered audio in video games. 'Interactive audio' consists of 'Sound events that react to the player's direct input' like footsteps and gunshots, whereas 'adaptive audio' is 'Sound that reacts to the game states' such as location, mood or health.[16] The former relates to precise gameplay, while the latter is not so directly affected by player activity. In terms of music in video games, while much works on a level of providing accompanying atmosphere in different locations, for example, and so is adaptive, some programming allows for developing music to be triggered by a succession of player inputs (such as proximity of enemies). Adaptive music is particularly effective in sophisticated action role-playing games with possible multiple paths. Good examples of this would include the later iterations of the Elder Scrolls series games, such as *Skyrim* (2011). A particularly complex form of synchronization comes from branching (horizontal resequencing) and layering (vertical remixing) music to fit momentary developments in the gameplay instigated by the player.[17] The process of pulling together disparate musical 'cues' involves direct joins, crossfades and masks, with a dedicated program controlling a database of music in the form of fragmentary loops, short transitions and longer musical pieces of varying lengths (see, for example, *Halo* (2001) and sequels). These sophisticated procedures yield constant variation, where music is integrated with the experience of the player. This means momentary change, synchronized precisely to events in the game, and this precise matching of musical development to action on screen owes something to the film tradition of 'Mickey-Mousing' but is far more sophisticated.[18] The relationship of image, player input and soundtrack is retained at a constantly close level, controlled by the programming and with momentary changes often not consciously perceived by the player.

Tradition has led to strong conventions in video game audio and in the relationship between sound and image. Some of these conventions are derived from other, earlier forms of audiovisual culture,[19] while some others are more specific to game design. Synchronization is fundamental for video

---

[15] Ibid., 36.

[16] Karen Collins, *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design* (Cambridge, MA: The MIT Press, 2008), 126.

[17] Tim van Geelen, 'Realising Groundbreaking Adaptive Music', in *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*, ed. Karen Collins (Aldershot: Ashgate, 2008), 93–102.

[18] Tim Summers, *Understanding Video Game Music* (Cambridge, UK: Cambridge University Press, 2016), 190.

[19] Ibid., 176.

games, but the relationship between sound and image can take appreciably different forms. It might be divided into four types: precise synchronization to gameplay, plesiochrony, forcing gameplay to fit music and asynchrony.

Player-Led Synchrony

Player-led synchronization has a succession of synch points where player input aligns precisely with sound and image. Player input can change screen activity, and this renders musical developments in line with the changes in game activity. This is what Collins calls 'interactive audio', and this principle is most evident in interactive sound effects where the game player provides some sort of input that immediately has an effect in the audiovisual output of the game. For instance, the pulling of a gun's trigger by the player requires a corresponding immediate gunshot sound, and requires a corresponding resultant action on screen where a target is hit (or not) by the bullet. This is a crucial process that provides an immersive effect, making the player believe on some level in the 'reality' of the gameplay and gameworld on screen. It correlates with our experience of the real world,[20] or at least provides a sense of a coherent world on screen even if it does not resemble our own. This triple lock holds together sound and image as an illusory unity but also holds the player in place as the most essential function. Indeed, the continued coherent immersive illusion of the game is held together by the intermittent appearance of such moments of direct, precise synchronization. The coherence of the experience is also aided by synchronized music, which forms a precise unity of visuals on screen, other sounds and gameplay activity. Music in such situations is dynamic, following player input to match location, mood and activity. This is *reactive* music that can change in a real-time mix in response to the action, depending directly on some degree of variable input from the player. It lacks the linear development of music as it is traditionally understood and indeed, each time a particular section of a game is played, the music might never be exactly the same.

An interesting case in point is *Dead Space* (2008), which has a particularly convoluted and intricate approach to its music. The game is set in the twenty-sixth century, when engineer Isaac has to fight his way through the mining space ship Ishimura that is filled with 'Necromorphs', who are the zombified remnants of its crew. To destroy them, Isaac has to dismember

---

[20] Alison McMahan, 'Immersion, Engagement and Presence: A Method for Analyzing 3-D Video Games', in *The Video Game Theory Reader*, ed. Bernard Perron and Mark J. P. Wolf (London: Routledge, 2003), 67–86, at 72.

them. Played in third person, *Dead Space* includes zero-gravity vacuum sections, and puzzle solving as well as combat. Jason Graves, the game's composer, approached the game events as a drama like a film. He stated: 'I always think of this the way I would have scored a film . . . but it's getting cut up into a giant puzzle and then reassembled in different ways depending on the game play.'[21] So, the aim is to follow the model of incidental music from film,[22] but in order to achieve this, the music needs to follow a complex procedure of real-time mixing. While some games only offer a repetitive music on/off experience, *Dead Space* offers a more sophisticated atmospheric and immersive musical soundtrack. Rather than simply branching, the game has four separate but related music streams playing all the time. These are 'creepy', 'tense', 'very tense' and 'chaotic'. The relationship between these parallel tracks is controlled by what the game designers in this case call 'fear emitters', which are potential dangers anchored in particular locations in the gameworld. The avatar's proximity to these beacons shapes and mixes those four streams, while dynamically altering relative volume and applying filters and other digital signal processing. This means that a constant variation of soundtrack is continually evident throughout *Dead Space*.[23] Rather than being organized like a traditional film score, primarily around musical themes, the music in *Dead Space* is built around the synchronization of musical development precisely to avatar activity, and thus player input in relation to game geography and gameplay.

Plesiochrony

Plesiochrony aims not to match all dynamic changes of gameplay, but instead to provide a general ambience, making a unity of sound and

---

[21] Ken McGorry, 'Scoring to Picture', in *Post Magazine*, November 2009, 39, accessed 15 October 2020, https://web.archive.org/web/20100126001756/http://www.jasongraves.com:80/press.

[22] Mark Sweeney notes that the game has two musical sound worlds: a neo-romantic one in cut scenes and a modernist one inspired by twentieth-century art music (of the sort used in horror films) which works for gameplay. The latter is reminiscent of Penderecki's music as used in *The Exorcist* (1973) and *The Shining* (1980). 'Isaac's Silence: Purposive Aesthetics in Dead Space', in *Ludomusicology: Approaches to Video Game Music*, ed. Michiel Kamp, Tim Summers and Mark Sweeney (Sheffield: Equinox, 2016), 172–97 at 190, 192.

[23] Don Veca, the game's audio director, created a scripting system he called 'Dead Script', which was on top of low-level audio drivers and middleware. An important aspect of this was what he called 'the creepy ambi patch', which was a grouping of sounds that constantly reappeared but in different forms, pitch-shifted, filtered and processed. These were also controlled by the 'fear emitters' but appeared more frequently when no action or notable events were happening. Paul Mac, 'Game Sound Special: Dead Space', in *Audio Media*, July 2009, 2–3.

image, and thus an immersive environment for gameplay. Player input is less important and might merely be at a level of triggering different music by moving to different locations. The term describes a general, imprecise synchronization.[24] Plesiochrony works in an isomorphic manner (as discussed earlier), matching together atmosphere, location and general mood. Music and image fuse together to make a 'whole', such as a unified environment, following the principles of being isomorphically related to atmosphere, location and mood. This might be characterized as a 'soft synchrony' and corresponds to Collins' notion of 'adaptive audio'. The music modifies with respect to gameplay in a broad sense, but does not change constantly in direct response to a succession of player inputs. The music in these cases becomes part of the environment, and becomes instituted in the player's mind as an emotionally charged and phenomenologically immersive experience. The music is often simply triggered, and plays on regardless of momentary gameplay. However, it nevertheless accomplishes an important role as a crucial part of the environment and atmosphere, indirectly guiding and affecting player activity. Indeed, perhaps its principal function is a general furnishing of 'environmental' aspects to the game, emphasizing mood, tone and atmosphere.[25] For instance, in *Quake*, the music at times quite crudely starts and stops, often with almost no interactive aspect. It is simply triggered by the player's avatar entering a new location. The score has something of the quality of diegetic ambience and at times could be taken to be the sound of the location. However, the sounds do not change when the avatar becomes immersed underwater, indicating that it is outside the game's diegesis.

While 3-D games like *Quake* followed a model evident in most first-person shooters, other 3-D games have adopted different approaches. Indie game *The Old City: Leviathan* (2015) is not based on skilful fighting action or thoughtful puzzling. It is a first-person 'walking game', where a detailed visual environment is open to the player's exploration. This game engages with a margin of video games history, games that are about phenomenological experience rather than progressive achievement and gameplay in the conventional sense. The music was a featured aspect of *The Old City: Leviathan*'s publicity, and the 'lack' of action-packed gameplay allows music to be foregrounded. As a counterpart to the visuals, the extensive and atmospheric music soundtrack is by Swedish dark ambient

---

[24] For more discussion of plesiochrony, see Donnelly, *Occult Aesthetics*, 181–3.

[25] Gernot Böhme points to atmosphere as a form of integrated, concrete relationship between human and environment. Gernot Böhme, *The Aesthetics of Atmospheres*, ed. Jean-Paul Thibaud (London: Routledge, 2017), 14.

industrial musician Atrium Carceri. The game's texture is emphasized by the player's slow movement around the city location in first person, which allows and encourages appreciation of the landscape. Indeed, the game developers were obsessed with images and sounds: on their promotional website, the discussion alights on the difficulty of visually rendering puddles. More generally, the Postmod website states: 'Players have the option to simply walk from start to finish, but the real meat of the game lies in the hidden nooks and crannies of the world; in secret areas, behind closed doors . . . '. The music is not only an integrated part of the experience, but also follows a similar process of being open to exploration and contemplation, as ambient music tends to be quite 'static' and lacks a sense of developmental movement. The fact that there is little real gameplay, apart from walking, gives music a remarkable position in the proceedings, what might be called 'front stage'. There is no need for dynamic music, and the music has the character of Atrium Carceri's other music, as atmospheric ambience.[26] The music is an equivalent to landscape in an isomorphic manner. It is like a continuum, develops slowly and has no startling changes. The aim is at enveloping ambience, with a vaguely solemn mood that matches the player's slow movement around the large, deserted cityscape. In a way, the game 'fulfils' the potential of the music, in that its character is 'programmatic' or ambient music. While the music appears somewhat indifferent to the player, an interactive relationship is less important here as the music is an integrated component of the game environment (where the world to a degree dominates the player), and music functions as triggered isomorphic atmosphere. Less concerned with gameplay and more directly concerned with embodying environment, it lacks notable dynamic shifts and development. This decentres the player and makes them a visitor in a large dominant and independent sound- and image-scape. However, there is a coherence to the game's world that synchronizes sound and image on a fundamental level.

Music-Led Asynchrony

In the first configuration of asynchrony, music *articulates* and sets controls for action and gameplay. Collins and others are less interested in this form of game audio, perhaps because initially it appears to be less defining and less specific to video games than 'dynamic audio'. However, I would

---

[26] Of course, this is importing wholesale music of a style that is not particular to video games, yet fits these particular games well.

suggest that asynchrony marks a significant tradition in game audio. With games that are timebound, music often can appear to set the time for gameplay. This is most clear, perhaps, with music-based games like *Guitar Hero* (2005) or *Dance Dance Revolution* (1998), where the fixed song length dictates the gameplay. A fine example are the games corralled as the *Wii Fit* (2007) programme, a massively popular series of exercises that mix the use of a dedicated balancing board with the console's movement sensor to appraise performance. The music for each task is functional, providing a basic atmosphere but also providing a sense of aesthetic structure to the often repetitive exercises.[27] Each piece is fairly banal and does not aim to detract from concentration, but provides some neutral but bright, positive-sounding wallpaper (all major keys, basic melodies, gentle rhythmic pulses and regular four-bar structures). In almost every case, the music articulates exercise activity. It is not 'interactive' as such, and the precise duration of each exercise is matched by the music. In other words, each piece of music was written to a precise timing, and then the game player has to work to that. They should not, or cannot, make any changes to the music, as might be the case in other video games. The *Wii Fit* has a number of exercise options available, each of which lasts a relatively small period of time. These are based on yoga, muscle exercises or synchronized movement. There are also some balancing activities, which are the closest challenges to those of traditional video games, and include ski-jumping and dance moves coordinated to precisely timed movements. The crucial point for the player is that they can tell when the music is coming to an end, and thus when the particular burst of exercise is about to finish. Repetition of the same exercises causes the player to know the music particularly well and to anticipate its conclusion. This is clearly a crucial psychological process, where the player is forced to synchronize themselves and their activities directly to the game, dictated by the music.

Similarly, in *Plants vs. Zombies* (2009), changes in the musical fabric are based on the musical structure and requirements, not the gameplay. It is a tower defence game in which the player defends a house using a lawn (in the initial version), where they plant anthropomorphic defensive vegetables to halt an onrush of cartoon zombies. Each level of the game has a different piece of accompanying music, which comprises electronic approximations of tuned percussion and other traditional instrument sounds. The relation of music to action is almost negligible, with the

---

[27] Composers include Toru Minegishi (who had worked on the *Legend of Zelda* games and *Super Mario 3D World*), Shiho Fujii and Manaka Tomigana.

regularity of the music furnishing a rigid and mechanical character to gameplay. The music simply progresses, almost in parallel to the unfolding of the game, as a homology to the relentless shambling movement of the zombies. The first levels take place during the day on the front lawn of the player's unseen house and are accompanied by a piece of music that simply loops. The recorded piece mechanically restarts at the point where it finishes, irrespective of the events in the game.[28] Because the music is not subject to interruptions from dynamic music systems, the music is free to feature a strong, regular rhythmic profile. The cue is based on a tango or habanera dance rhythm. The very regularity of dances means that, as an accompaniment to audiovisual culture, they tend to marshal the proceedings, to make the action feel like it is moving to the beat of the dance rather than following any external logic. In *Plants vs. Zombies*, this effect is compounded by the totally regular structure of the music (based on four-bar units, with successive melodies featuring particular instruments: oboe, strings and pizzicato strings respectively).[29] There is a clear sense of integrity from the music's regular rhythmic structure in the face of variable timings in the gameplay. The harmony never strays too far from the key of A minor (despite an F 7th chord), and the slow tango rhythm is held in the bass line, which plays chord tones with the occasional short chromatic run between pitches. However, the continuity of this music is halted by the player's successful negotiation of the level, which triggers a burst of jazz guitar to crudely blot out the existing music. *Plants vs. Zombies*' music conceivably could fit another game with a profoundly different character. Having noted this, the cue's aptness might be connected most directly with the game on a deeper level, where the music's regularity relates to the unceasing regularity of the gameplay. The cue is not synchronized to the action, apart from the concluding segment of the level. In this, a drum beat enters as an accompaniment to the existing music, appearing kinetically to choreograph movement through grabbing proceedings by the scruff of the neck, as what is billed on screen as a 'massive wave of zombies' approaches at the level's conclusion. The assumption is that the beat matches the excitement of the action (and chaotic simultaneity on screen). However,

---

[28] Although the player's actions trigger musically derived sounds, forming something of a random sound 'soloing' performed over the top of the musical bed, it is difficult to conceive this as a coherent piece of music.

[29] The opening tango section comprises sixteen bars, followed by eight bars of oboe melody (the same four repeated), then an orphan four-bar drop-out section leading to pizzicato strings for eight bars, followed by the same section repeated, with added sustained strings for eight bars, and finally, a section of piano arpeggios of sixteen bars (the same four repeated four times), after which the whole piece simply repeats.

again, if we turn the sound off, it does not have a significant impact on the experience of the game and arguably none at all on the gameplay. In summary, the time of the music matches the game section's pre-existing structure, and the player has to work to this temporal and dynamic agenda rather than control it.

Parallel-Path Asynchrony

Unsynchronized music can also have a relationship of indifference to the game, and carry on irrespective of its action. It might simply co-exist with it, as a presence of ambiguous substance, and might easily be removed without significantly impairing the experience of the game. This situation is relatively common for mobile and iOS games and other games where sound is unimportant. Here, the music is not integrated strongly into the player's experience. This relationship of asynchrony between the music and gameplay is embodied by this form of 'non-functional' music that adds little or nothing to the game and might easily be removed or replaced. Such music often has its own integrity as a recording and can finish abruptly, when the player concludes a game section (either successfully or unsuccessfully), sometimes interrupted by another musical passage. This owes something to the tradition of earlier arcade games, and although it may seem crude in comparison with the processes of dynamic music, it is nevertheless an effective phenomenon and thus persists. The way this non-interactive music carries on irrespective of gameplay makes it correspond to so-called 'anempathetic' music, where music seems indifferent to the on-screen action. Anempathetic music has been theorized in relation to film, but clearly has a relevance in all audiovisual culture. Michel Chion discusses such situations, which have the potential to 'short-circuit' simple emotional congruence to replace it with a heightened state of emotional confusion.[30]

Some games simply trigger and loop their music. Although 'interactive dynamic music' can be a remarkable component of video games and is evident in many so-called 'four-star' prestige video games, in its sophisticated form, it is hardly the dominant form of video game music. Disconnected 'non-dynamic' music exemplifies a strong tradition in game music, which runs back to the arcade.[31] While some might imagine this is a retrogressive format, and that arcade games were simplistic and aesthetically unsophisticated, this earliest form of video game sound and

---

[30] Chion, *Audio-Vision*, 8–9.
[31] Donnelly, 'Lawn of the Dead', 154.

music remains a highly functional option for many contemporary games. For instance, this is highly evident in iOS and other mobile games, and games which lack dynamics and/or utilize highly restricted spatial schemes, such as simple puzzle games or games with heavily constrained gameplay. A good example is action platformer *Crash Bandicoot* (1996), which has a sonic backdrop of kinetic music emphasizing synthesized tuned percussion, while earcons/auditory icons and sound effects for game events occupy the sonic foreground. The music changes with each level, but has no requirement to enter a sophisticated process of adaptation to screen activity. Yet in purely sonic terms, the game achieves a complex interactive musical melange of the game score in the background with the highly musical sounds triggered by the avatar's activities in the gameplay. Yet the game can happily be played without sound. The musical background can add to the excitement of playing, evident in arcade games such as jet-ski simulator *Aqua Jet* (1996), which required the player to 'qualify' through completing sections in an allotted time. The music merely formed an indistinct but energetic sonic wall behind the loud sound effects. Similarly, in driving game *Crazy Taxi* (1999), 'game time' ticks down and then is replenished if the player is successful. Here, songs (including punk rock by The Offspring) keep going, irrespective of action. The music can chop and change when a section is finished, and when one song finishes another starts. The situation bears resemblance to the Grand Theft Auto series (1997–2013) with its radio station of songs that are not synchronized with action.

Candy Crush Saga (2012) is an extremely successful mobile game of the 'match three' puzzle variety. Music is less than essential for successful gameplay, although the game's music is highly effective despite many people playing the game 'silent'. The merest repetition of the music can convince the player of its qualities.[32] It is simply a looped recording and in no way dynamic in relation to the game. The music has a fairly basic character, with a swinging 12/8 rhythm and clearly articulated chord changes (I-IV-IVm-I-V) at regular intervals. However, one notable aspect of the music is that at times it is extremely *rubato*, speeding up and slowing down, which helps give a sense that the music is not regimented, and the player indeed can take variable time with making a move. The player's successive moves can sometimes be very rapid and sometimes taken at their leisure. Whatever happens, the music carries on regardless. However,

---

[32] The 'repetition effect' tests music's durability, although the cumulative effect of repetition can persuade a listener. I hated the music at first, but after playing the game, ended up appreciating it.

it is intriguing that the music contains a moment of drama that pulls slightly outside of the banal and predictable framework. The shift from major to minor chord on the same root note is a surprising flourish in a continuum of highly predictable music. This exists perhaps to make the music more interesting in itself, yet this (slightly) dramatic change signalled in the music is *not* tied to gameplay at all. It is subject to the random relationship between music and gameplay: the music can suggest some drama where there is none in the game. It might be argued that this applies a moment of psychological pressure to the player, who might be taking time over their next move, or it may not. However, while the music is a repeated recording with no connection to the gameplay, oddly it has a less mechanical character than some game music. The *rubato* performance and jazzy swing of the music can appear less strictly regimented than some dynamic music, which needs to deal in precise beats and quantified rhythms to hold together its disparate stem elements.[33]

Conclusion

The process of 'triggering' is the key to the whole process of synchronization in video games and may be noticeable for the player or not. In some cases, the trigger may simply inaugurate more variations in the musical accompaniment, but in others it can cue the beginning of a whole new piece of music. A certain activity embarked upon by the player triggers a change in the game audio. This locks the synchronization of the player's activity with sound and image elements in the game. Indeed, this might be taken as one of the defining aspects of video games more generally. This interactive point functions as a switch where the player triggers sound activity, such as when moving an avatar to a new location triggers changes in the music and ambient sound.[34] This physical, technological fact manifests the heart of the sound-and-image synch that holds video games together. However, we should remember that rather than simply being a technical procedure, crucially it is also a psychological one. It initiates and signals a significant change in the gameplay (more tense gameplay, the start of a new section of play, etc.).

---

[33] 'Stems' are musical parts that work through fitting intimately together rather than on their own. Arguably, the process of using stems comes from digital musical culture, where it is easy to group together recorded music channels in so-called 'submixes'.

[34] Sweet notes that these points are called 'hooks', which describe the commands sent from player input, through the game engine to the audio, so named for the game 'hooking into' the music engine. Sweet, *Writing Interactive Music*, 28.

There has been much writing by scholars and theorists of video game music about interactive audio.[35] However, 'interactivity' presents severe limits as a theoretical tool and analytical concept. Similarly, an analytical concept imported from film analysis used for dealing with video game audio – the distinction between diegetic and non-diegetic – also proves limited in its relevance and ability as a means of analysis. Instead of these two concepts, a division of music through attending to synchronization might be more fruitful for video game analysis. This would register the essential similarities between, for example, non-diegetic incidental music and diegetic sound environment for a particular location. Similarly, the dynamic changes that might take place in an interactive musical score could also follow a similar procedure in diegetic sound effects, when a player moves an avatar around a particular on-screen location. Analytical distinctions of music in video games should not be determined purely by mode of production. This has led to a focus on interactive video game music, which may well not be phenomenologically perceived by the player. This risks a focus on the underlying production, the coding level, to the detriment of addressing the phenomenological experience of the surface of gameplay.[36] The electro-mechanical perfection of time at the centre of modern culture endures at the heart of video games, whose mechanistic structures work invisibly to instil a specific psychological state in the player.

---

[35] Not only has Karen Collins written about different forms of interactive audio, but also Michael Liebe, who notes that there might be three broad categories of music interaction: 'Linear' (which cannot be changed by the player), 'Reactive' (which is triggered by player actions) and 'Proactive' (where the player must follow the game). Michael Liebe, 'Interactivity and Music in Computer Games', in *Music and Game: Perspectives on a Popular Alliance*, ed. Peter Moormann (Wiesbaden: Springer, 2013), 41–62, at 47–8.

[36] Frans Mäyrä notes that players, ' . . . access both a "shell" (representational layers) as well as the "core" (the gameplay).' Mäyrä, 'Getting into the Game,' 317.