# Accepted manuscript

## Article 2 – "A Swarm of Sound": Audiovisual Immersion in Björk's VR Video *Family*

*Zack Bresler and Stan Hawkins*

## Introduction

Technological advances in pop music video productions have undergone significant changes in recent years, with performances increasingly spectacularized through the aid of new generations of camera devices and editing software. The advent of the internet has altered modes of consumption, sharing, and dissemination of the music video through the portals of Facebook, Spotify, Instagram, TikTok, Twitter, and YouTube (see Richardson, Gorbman & Vernallis 2013; Vernallis 2013; Korsgaard 2013, 2017; Burns and Hawkins 2019, p. 2). Music videos provide recourse for evaluating representations in new media technologies in a bid to understand the dynamic workings of artists from a range of disciplinary perspectives. For the purpose of this article, questions are raised that deal with aspects of immersion and interactive engagement. How do properties of compositional design in the virtual reality (VR)[15] music video function in establishing notions of space? What is the listener's role in multidimensional spatial environments? And how do the congruences of audiovisual sensory data enhance the causes and effects of sonic immersion?

Our starting point is to define audiovisual immersion as a pleasurable state of consciousness that is characterized by complete absorption, a result of the dialectic interactions between a viewing subject and compelling audiovisual experience. We have provided a model of the virtual audiovisual space (VAVS) that has as its objective to conceptualize experiences of audiovisual immersion in music. By extending ideas on 'virtual acoustic space' (Wishart 1996), this model emphasizes the relationship of visual imagery to sound and how this enhances agency, both on the part of the performer and viewer. Our goal is to consider sound and image

---

[15] Virtual reality (VR) describes a computer-generated, three-dimensional environment that can be experienced, explored and/or interacted with through the use of VR peripherals, such as isolating visual 3D headsets. Examples of VR systems include the Oculus Rift, HTC Vive, and Valve Index.

together as focal points for analysis in our understanding of virtual reality in music videos.[16]

In our study of the VR video, 'Family' (2019)[17] by the Icelandic pop icon Björk, we have considered the intermeshing audiovisual signifiers within a soundscape that enhances the sensations of immersion. Björk's progressive approach to technologies of immersion and interactivity has prompted similar scholarship, as evident in the studies of the "app album" *Biophilia* by Nicola Dibben (2013). Some of Dibben's claims about the app album can be applied to *Vulnicura VR*:

> *Biophilia* (re)introduced multimodality into digital audiovisual formats and used this to realize a creative vision of intuitive and embodied forms of music making and learning in which the natural word provides productive metaphors for emotional experiences and musical processes (2013, p. 699).

Working out the conditions of the multimodal virtual space in 'Family', we choose to concentrate primarily on the aesthetic effects of the VR performance. In the main, our model functions as a platform for considering the experiences of space and temporality within a highly active VR context; a context that functions as a staged environment that implies different things to every listener, culturally and socially. For instance, the design of sound and imagery might be perceived as surrealistic in one context yet entirely different in another. Hence, aesthetic experiences are contingent on a range of factors, and the analytical insights we provide are predicated upon personal interpretations and textual analyses.

We would suggest that the experience of audiovisual entertainment positions sound and image in the listener's memory. Lelio Camilleri's model of the sonic space (2010, p. 202) addresses this in terms of the 'localised space', the spectral space' and the morphological space', to which we add the 'aesthetic space', where sound and image synthesize in the audiovisual sensory experience. Such spatial

---

[16] Congruent with this goal, Anders Aktor Liljedahl has drawn attention to the way that studies of audiovisual media, including music video generally prioritise the visual and silence the music (2019, pp. 166-167).

[17] 'Family' was first released on the 2015 album *Vulnicura,* accompanied by 'moving album cover' featuring a short version of the song. The first VR video for 'Family' premiered in November 2016 at Harpa in Reykjavik  (https://grapevine.is/icelandic-culture/art/2016/11/02/bjork-digital-opens-today/). The version discussed in this article is the video as it was re-formatted and re-mastered for consumer VR devices and released on the digital album *Vulnicura VR* in September 2019 on the Steam PC gaming platform.

combinations can be comprehended in terms of the aesthetics of sensory perception, creating the feeling of saturation. Integral to space in the music video are the properties of compositional design and the materiality of numerous 'stylistic and technical' codes (Hawkins 2002, pp. 9-12). Thus, audiovisual space accommodates the features and effects of sensory perception that instate the dramaturgy of a VR video performance. Given that music videos are contextualised within a mediascape, we also consider how *intermediality*, as defined by references, evocations, and techniques, impacts on VR productions.[18] One might posit that music videos are audiovisual compositional designs in themselves, their combined features mediated across any number of platforms during a performance. This would imply then that intermediality enables listeners to engage actively with the structural features of design.

In devising our VAVS model, we have been keen to highlight the attributes of "source bonding",[19] or the connection between heard sounds and their supposed causes (Smalley 1997, p. 110) that emanate also from shared experiences as they unfold through time and sensations of immersion. In the VR video we consider, it is as much the technological staging of space (in terms of texture, temporality, and gesture) as the musical features (rhythm, harmony, and melody) that define the 'aesthetic space'. Björk's VR performance comes across mysterious, if not scintillating since she breaks with many of the norms and traditions of the standard pop video format, which arguably becomes a metaphor for severing the constrictions of the conventional family unit. Our position is that 'Family' is derived from a pool of spatialities that denote a new audiovisual compositional domain and that enables the music to reach the viewer in a powerful and visceral way; that music immerses us within VR imagery is a highly personal affair.

## The Virtual Audiovisual Space (VAVS)

If spatialities are integral to audiovisual contexts, being immersed in a music video is intermedial and multi-faceted. In Jem Kelly's words, the music video is "already a hybrid medium, comprising audio and visual forms and structures that intersect

[18] Intermediality as a term originated from intertextuality in 1983, which spawned a movement of intermediality studies led by German scholars, Aage Hansen-Löwe, Claus Clüver, Irina Rajewsky and Werner Wolf.
[19] Smalley's notion of 'bonding' goes beyond the idea of source-bonding used in this article. For Smalley, bonding simply is the way that sound and context are related, and source-bonding is one mode of this relation.

and interrelate in ways that can be described as intermedial" (2019, p. 220). We also adhere to the concept of 'multimodality', as theorised by Lori Burns, who considers "multimodality to comprise the artistic integration of multiple semiotic modes within one media text" (2018, p. 96). Part of what constitutes perceptions of agency is at the centre of the listening and viewing experience, and the video offers a glimpse of a specific context through a multimodal composite. The visualisation of the performance through immersion propels the viewer into a different interpretive space. Björk's VR video accomplishes this through the intermedial and multimodal relations of the sonic and the visual,20 which forms the basis of the audiovisual compositional design in the context of immersion in VR music video experience. Accordingly, we focus primarily on the context in which there is a "transgression of boundaries between what is conventionally perceived as distinct media" (Wolf 2015, p. 461).

Trevor Wishart's (1996) concept of virtual acoustic space (VAS) provides an in-depth insight into the compositional design of electroacoustic music and serves as an inspiration for our model of virtual audiovisual space (VAVS). In addition to Wishart's VAS concept, our study takes heed of Camilleri's sonic space (2010) and Denis Smalley's notion source-bonded spaces (2007, p. 38). In particular, we identify four dimensions in the audiovisual space: (1) sonic spaces:21 the environments in which sonic objects are placed, and their morphologies; (2) visual spaces: the virtual immersive imagery that constitutes what is visible (3) source-bonded spaces: the spaces in which the listening agent connects those objects to meanings through experience; and (4) aesthetic spaces: the abstract spaces where sound and image combine in the listener's memory to create meaning that transcends the source-bonding connections between the two (see Figure 1).

---

[20] We are acutely aware of the smudging of conceptual lines between intermediality and multimodality in this discussion so far, and therefore acknowledge the fact that all media comprise 'mixed media'. In our understanding, intermediality is the relationship between two media, such as music and imagery, and how they reference one another, while multimodality pertains to the application of variable literacies within one medium. For example, a music video performance involves the comprehension of language, culture, politics, and geography.

[21] Camilleri's model of 'sonic space' addresses the "space in which the [acousmatic] piece unfolds" (2010, p. 201). This three-dimensional model, which consists of localised space (the "space into which sounds are placed"), spectral space (the sensory understanding of timbre and disposition), and morphological space (the temporal aspect of space), accounts for the placement and disposition of sound objects as well as their morphologies, the ways in which such sounds are perceived temporally (ibid., p. 202).

In our model, the concept of sonic space pertains to the space created by auditory events and the way these events change in time and space. Indeed, the visual space denotes an extension of sonic space that accounts for the additional sense of vision. Just as sound can be described in terms of the positioning, disposition, and temporal unfolding of sonic objects, so can the visual space be understood in the same terms for visual objects. From this it is apparent that the sonic and visual spaces are not experienced or understood as distinct, although we argue it is productive to interpret them independently.[22]
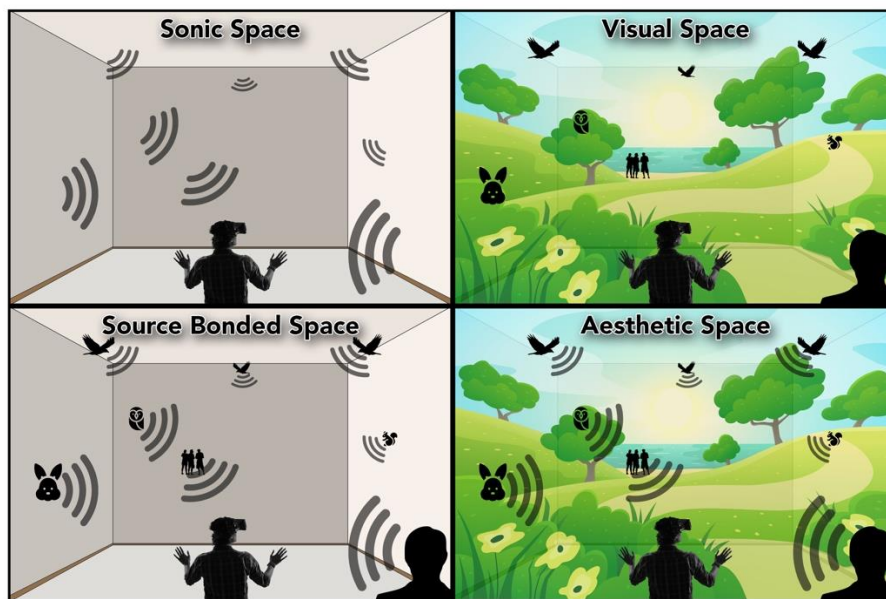


Figure 1: Virtual Audiovisual Space (VAVS)

Smalley has identified the spatialities created in the relationship between sound causes and their imagined sources as 'source-bonded spaces' (2007, p 38).[23] Accordingly, our model implements this to illuminate the ways in which the listener comprehends sonic and visual presentations via these assumptions.[24] As a

---

[22] For a comprehensive overview of the discourses around inter- and transmediality in music studies, see Werner Wolf's chapter "Literature and Music: Theory" from the 2015 De Gruyter *Handbook of Intermediality: Literature – Image – Sound – Music*. While our emphasis is more on intermediality, we have also considered that the music video might be framed within a *transmedial* context, where the media "unfolds across multiple media platforms, with each new text making a distinctive and valuable contribution to the whole" (Jenkins, 2006, pp. 95-96).

[23] Dwelling on Smalley's notion of 'source-bonding', as sounds are encountered they are processed in terms of their assumed causes, regardless of the fact that the formalization of the sound through recording, processing, and re-production through speakers has distorted the sonic image (1997, p. 110). Arguably, this has a corollary in Michel Chion's notion of 'causal listening' in the audiovisual experience (1994, pp. 25-28).

[24] As Smalley points out, this conceptualization of space is resonant with Lefebvre's notion of space as a social morphology. See Lefebvre's *The production of space,* originally published in 1974 and translated into English in 1991 by Donald Nicholson-Smith. Here, we relate this to acousmatic sound to suggest that

facet of immersive media, the 'source-bonded space' brings to the fore the listener's own interpretive perspective, and hence imports their agency into the model. In turn, this raises the idea of the artist's and listener's interaction during immersion, a significant aesthetic and experiential entity.

Given that the aesthetic space comprises a zone in which the multimodal experience is cognized and synthesized, it contains different modes of source-bonding that engage the viewer in ways that transcend either the sonic or the visual. In this sense, the aesthetic space is defined by the connections between sound and image that do not rely on exact sound-to-image source-bonding be understood. From this it is apparent that a higher-level of synthesis occurs simultaneously to the independent modes of sonic and visual understanding. Holly Rogers has suggested that the "audio basis, together with its continual motion, posits for the video image an existence in the musical sphere and vice versa… and its meaning no longer needs to be 'emergent' as it materializes, unified, at the moment of its creation" (2011, p. 410). Thus, in the context of immersive VR music video, aesthetic spaces signify a type of meta-spatiality, where new modes of sonic meaning arise (in contrast to the music without the video).

While source-bonded space relies on interpretation, it does so in the ecological rather than the hermeneutic sense, which, in our model, accommodates the domain of aesthetic space. This implies that interpretation in the source-bonded space does not necessarily require cognition, as this is done pre-consciously and corresponds to the viewer simply understanding the 'cause' of a sound and its relational context in the music. Eric Clarke has insisted that in this ecological mode of perception, "to hear a sound and recognize what it is… is to understand its perceptual meaning (2005, p. 7). Moreover, understanding the aesthetic space is a hermeneutic project. Accordingly, we adhere to Lawrence Kramer's call for "open interpretation", which "aims not to reproduce its premises but to produce something from them" (2011, p. 2). In this sense, source-bonded spaces represent the literal connections between sound and image, while the aesthetic space corresponds to the metaphors and meanings that we create in interpreting their connection.

Granted, none of these spatialities are mutually exclusive, implying that audiovisual cognition is only possible within the temporal framework of recalling

---

sounds produce the space they occupy through their understood relational attitude towards and with the listener.

the composition retrospectively as a singular event. As a result, notions of temporality are central to the audiovisual experience. Bearing this in mind, various strata are analysed that are deemed pertinent for inspecting the properties of immersion.

## Immersion and Compositional Design

Returning to our concept of audiovisual immersion we now consider its inextricable ties to compositional design. Anders Aktor Liljedahl has stated that immersive and interactive music videos "suggest both an infinite set of outcomes and an enclosed range of possibilities," (2019, p. 183). To this we might add the immersive music experience, which creates endless possibilities of relating to compositional design. From one perspective, the 3D audiovisual experience 'de-idealizes' viewers as they become in themselves non-static and dynamic objects of the compositional design—a part of the music that experiences and creates meanings through interaction with the stage. Immersing oneself in surround sound and 3D music imagery entails grasping music production aesthetics more broadly. Just as the size and shape of the stage certainly matters, it is only a frame for accommodating the normative structures that define pop productions. As such, features of staging, when applied to and analysed in different formats, provide us with insights into the complexities of audiovisual immersion.

### *Quagmires of Immersion*

Activating the term "immersion" necessitates closer inspection. Two critical points arise: first, "immersive audio" and "immersive media" have in recent years become buzzwords, largely used for marketing speakers, televisions, gaming systems, VR headsets, mobile phones, and any number of commercial electronics. In this sense, the term refers mainly to a type of media format.[25] In a broader audiovisual multimedia context, 'immersive media' primarily refers to virtual and augmented reality, accessible with devices such as the Oculus Rift or HTC Vive VR headsets. In studying VR and its immersive effects, Mel Slater has identified

---

[25] An example of an immersive audio format is Dolby Atmos, an object-based format that was developed for use in cinema and which has recently come into use for 3D music streaming on the HD tier of Amazon's Prime Music service. Atmos uses a standard surround sound configuration with an additional surround layer positioned a distance above the listening position. https://news.dolby.com/en-WW/182472-fall-in-love-with-music-all-over-again-with-dolby-atmos-on-echo-studio-and-amazon-music-hd

that immersive systems can be typologized based on their degree of immersive effect, characterizing virtual reality systems by their set of valid actions, that is, "the actions that a participant can take that can result in changes in perception or changes to the environment" (Slater 2009, p. 3550). Here, immersion is defined as "a property of the valid actions that are possible within the system" (ibid., p. 3551), and systems with more types and/or better qualities of valid actions are considered more immersive.

The second critical point is that, while Slater's notion of immersion is derived technologically, it is clear that immersion is also phenomenological. For example, it can be considered in terms of absorption, the state of consciousness that Graham Jamieson defined as "an effortless, non-volitional quality of deep involvement with the objects of consciousness" (2005, p. 120) and which is contrasted against an instrumental disposition which requires serious cognitive effort and planning.[26] Ruth Herbert goes further than Jamieson by applying the notion of absorption to the experiences of music listening, suggesting that "absorption and dissociation are best understood as processes that are subsumed within trance" (2011, p. 85). For the purposes of our definition of audiovisual immersion, we concur with Herbert's definition of "absorbed trancing", "characterized by imaginative involvement" that arises from "apparently passive yet still creative involvements such as listening to stories, listening to music, daydreaming, reading and imagining fiction, plus circumstances such as travelling on a train or being in a crowded place" (ibid., p. 134). Thus, retrieving an experience of being immersed need not be an audio or visual experience at all; it can be one of experiencing one's favourite music in any format, and it can just as well be another activity such as reading a book or taking a walk. The VR music video, however, combines many elements of absorption and trancing, enabling heightened sensory experiences that can lead to audiovisual immersion.

### *Immersion and Agency*

Worth considering are the ways in which immersive media may more easily facilitate the immersive experience. One way that this occurs is through increased

---

[26] Immersion has been correlated to states of flow (Csikszentmihalyi 1990); immersive experiences are recalled in moments when slowing down allows for periods of intense focus. However, flow has been problematised in relation to immersion since it is "an extreme experience where goals, challenge, and skill converge. As such, flow is an all or nothing experience" (Sanders & Cairns 2010, p. 161).

agency on the part of the viewer. When watching a music video on a 2D surface, such as a laptop screen, the viewer assumes a passive role; engagement might well feel like interaction although the viewer is not staged in the same way as when entering a VR experience. Effectively, the 3D visual experience of VR engulfs the viewer. As with the Björk video 'Family', we have noted that this requires multiple viewings, since the decision to focus on one particular entity will inevitably lead to missing out on another. Notably, René Idrovo and Sandra Pauletto have extended ideas found in the work by Michel Chion (1994, pp. 90-91) and Rick Altman (1992, p. 60) on diegetic perspective in film sound, terming the "immersive point of audition" as "a sound design approach that aims to locate the audience on a specific point within the diegesis, and thus lures us to be transported into the story by providing an immersive representation of sound" (2019, p. 39). Extending this further, we would suggest that the agency of the viewer in a VR context places them *within* an audiovisual scene, signalling an *interactive* point of audition, whereby the viewer is not only placed on a point within the diegesis, but also has control over its perspectival transformations.

In the current research on user experience in media and games, concepts of immersion and engagement are critical to understanding the role of music. Engagement has been defined as our "ability to recognize a work's overturning or conjoining conflicting schemas from a perspective outside the text" and immersion as being "completely absorbed within the ebb and flow of schema" (Douglas & Hargadon 2000, p. 154). In video games, engagement through interactivity is a fundamental aspect of immersion and is typically considered in an embodied way, wherein "game controllers can become an extension of the body into the virtual world" (Collins 2013, p. 41). On the other hand, narratological approaches have often seen high degrees of agency as being antecedent to immersive experience, since it breaks the story into small, difficult to synthesize portions, while large complex stories require the rigidity of fixed, non-agential story structure (Douglas & Hargadon 2000, p. 155). We would argue that these seemingly contradictory notions of immersion are simply different classes of experience which constitute different modes of trancing. In general, audiovisual immersion in VR music videos is more like that in video games, where agency through interactivity is key, and where "the ability to move through virtual landscapes can be pleasurable in itself" (Murray 2016, p. 125).

In the *Vulnicura VR* music videos, as in many virtual reality experiences, one literally has the sense of being 'spaced out' through sheer immersion.[27] This results in temporary notions of de-virtuality, bridging the phenomenological gap between sensations of the real and the virtual. This relates to Jay David Bolter and Richard Grusin's notion that "virtual reality has become a cultural metaphor for the ideal of perfect mediation" (2000, p. 161). That is to say, that through the intensity of its means of mediation, it carries the potential to dissolve the very feeling of mediation. Complicating the boundaries between virtuality and reality is the notion of the digital, which promises that "our creative thoughts and imagination (i.e., the virtual) can be either transformed or *nearly* transformed into reality and actuality through digital means" (Rambarran 2021, p. 1, emphasis in original). As such, a major part of what constitutes this transformation in VR occurs through an interactive relation between taking "meaningful action and see[ing] the results of our decisions and choices (Murray 2016, p. 123). Again, the impact of agency on experiences of audiovisual immersion is critical.

In considering agency and immersion, we wish to stress the distinction between immersion and interactivity. While interactivity might be part of immersive media, it is not necessarily a part of immersive experience. Hitherto we have described audiovisual immersion as that sense of absorption within the media experience. Alternatively, interactivity is reserved for those instances in which the listener/viewer becomes an active creative agent. In addressing interactive installation art, Rogers states, "sound and image can be manipulated by visitors in order to create individual audiovisual pathways; or visitors in different location can be drawn together via technological intervention" (2014, p. 8). This suggests that a continuum of agency is possible within music and media, where at one extreme the listener is presented with a media at a distance, and at the other extreme they are transported into an interactive sound-world as a freely creative agent. Stereo music and 2D video are, in most contexts, closer to the former, while virtual reality is closer to the latter. However, as previously intimated, there are a number of other factors that contribute to the phenomenology of immersion, including the ability to engage meaningfully with the presented content in an extra-textual way. Accordingly, features of immersive media, such as surround and 3D

---

[27] Discourses on immersion and music listening are numerous. Some supplementary texts are worth mentioning here, including Tia DeNora's *Music in Everyday Life* (2000), Joel Krueger's article "Enacting Music Experience" (2009) and Simon Høffding's *A phenomenology of miusical abosorbtion* (2019).

sound and imagery, freedom of movement and position of the listener, and degrees of interactivity, create extra possibilities of immersive experiences, especially when content and context is made meaningful for the recipient.

## *Compositional design and perceptions of listening*

Features of compositional design – a conceptual framework for describing how musical codes coalesce within a sound environment – lead to a holistic understanding of a track. Stylistic and technical codes can be utilized as part of a hermeneutic approach to music analysis (Hawkins 2002, p. 10-12), where the 'pop score' invariably comprises musical, social, and cultural objects that are coded and contextualized in such a way that the listener comprehends them as sonic representations of physical spaces and places. In addition, metaphorical, social, and cultural phenomena impact on our perceptions of compositional design and its structures.

Given that the central analytical framework for understanding the pop score is through its 'sound', then the use of the 3D sound stage in VR video has a significant impact. As the sound stage surrounds the listener, new subject positions, modes of performance, and proxemic relationships between performer and audience emerge (Bresler 2021). As we have highlighted in our analysis, this is often due to the simplicity of certain sounds appearing to emanate from unexpected locations, matching (or not) their visual counterparts in ways that push and pull the viewer's attention in multiple directions. In other cases, this is created by staging reverb, delay, and other secondary music processing behind the listener to create the feeling of particular acoustic spaces and places, or to literally surround the listener in a sea of voices or textures.
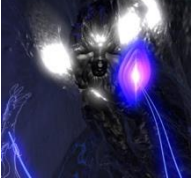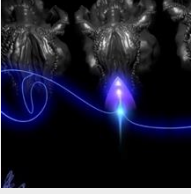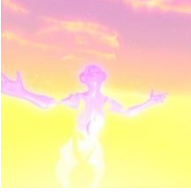
In applying notions of compositional design to the staging of audiovisual immersion, we are compelled to ask: where and how is the listener situated? In traditional media, the listener can often be perceived as metaphorically staged in an audience position with respect to the performance. However, this idea begins to disintegrate when creative and spatial formats 'surround' and engulf the listener. In immersive and interactive media, the stage is shared in an active way, with the listener positioned as a 'staged object' of the compositional design, their presence implying agential self-positioning. Certainly, the boundary between 'traditional' and 'immersive' media are not that distinct, and we do not intend to imply that films, stereo music, television, or any other form of media is incapable of creating such immersive experiences and staged subject positions. However, it is clear that

staging in VR is ontologically different from film with surround sound, for example, since the viewer expresses additional agency not only through their placement on the stage, but through their active participation in their own point of audition.

To demonstrate this we have undertaken an analysis of Björk's VR music video *Family*, released on the album *Vulnicura Virtual Reality* (2019).[28] Figure 2 provides a structural overview of the video and track through eleven discrete sections, with a focus on visual details, audio design, and the overall immersive effects (as we the authors encounter them). This table represents a semiotic analysis of the track, and functions as an aid to understanding the audiovisual processes and elements inherent in the VR music video.

Figure 2: 'Family' from Björk's Vulnicura VR, a detailed close reading (Next 2 pages)

---

[28] The researchers viewed the video on an Oculus Rift VR headset. At the time of writing, readers who are interested in viewing the video or the complete *Vulnicura VR* album will require this or a similar headset (such as a Valve Index or HTC Vive, and a PC computer with a suitable graphics processor. The album is, at this time, only available for purchase on the Steam PC gaming platform.

| Screenshot / Time | Visual Details | Audio Design | Immersive Effects |
|---|---|---|---|
| **(0:00 – 0:22)**  | The video commences in the dark, with a mysterious, purple, oblate luminous object. The viewer's digitized hands become visible as glimpses of flashing light reveal the environment as a tunnel-like cave, in which the viewer is moving with the object slowly forward. | The audio track begins with low strings uneasily blending between two notes a major 2nd apart. Very quickly we hear a loud, low-frequency impulse, followed by dissonant electronic sounds that resemble feedback and digital stutter. The impulses repeat regularly. | The first sensation is one of darkness, the appearance of a purple object, and sighting one's digital hands. The movement of exploration at the outset serves to establish the sound as immersive. It is immediately apparent to the viewer that both the visual and auditory aspects can be controlled through movement. |
| **(0:23 – 1:36)**  | Björk, represented as a digitized body, flashes in and out—a translucent figure with child-like buns in her hair. We move with her slowly through the cave. As the flashing lights continue, tentacle-like structures appear behind the oblate object, out of which purple streamers begin to pour. | The sounds already established continue to along the same lines as Björk begins to sing. She starts with the line, "Is there a place, where I can pay respects, for the death of my family…" Continuing, the voice introduces a kind of call-and-response with Björk's voice replicated and panned to the rear of the soundstage. | At this point, the viewer will be aware that that their hands can be made to move in a kind of slow conducting pattern if the trigger is pressed on the controller. This also causes streamers to pour from the object to track the hands. |
| **(1:37 – 2:16)**  | Björk, in sync with a lyrical cue, falls to her knees in front of the viewer. | The singer laments, "So where do I go, to make an offering? I fall on my knees." By now, the intensity and volume of the strings slowly increase in the background. | As Björk sings, vocal directionality is towards her physical manifestation. The contrapuntal lines of the voice, however, are panned away from the front, immersing the viewer completely. The voices to the rear seem at times distant as well, drawing the focus forward in the direction of movement. |
| **(2:17 – 3:05)**  | At this point, the end of the cave comes into sight, as the music builds slowly towards a climax. The end of this section is as visible as it is audible. In the distance a sculpture apparition appears, albeit difficult to discern. | Breaking from the previous call-and-response structure, the two vocal lines merge together contrapuntally, as Björk sings "So where do I go, to make an offering, to mourn our miraculous triangle, Father, mother, child..." with each line, the number of harmony and counterpoint voices increases. | Entry of new vocal harmonies provide a sense of lateral imaging, as the lead vocal now assumes more space in the front. New vocal strands surround the listener, some of which are distant cries, others like whispers in the ear. Visually, tentacle structures now surround the viewer, constructing a kind of magical mothership that propels us through the cave. |
| **(3:06 – 3:49)**  | The object in the form of a wound becomes more monotone and simplistic, yet still present alongside the translucent body figure. Surrounding the listener are several black and grey sculptures of Björk, bending backwards with her hands touching her feet, and rolling in that direction out of sync with one another. | At this point there is a dramatic change in the music in the form of a transition passage. The strings now play in an erratic, pizzicato Penderecki-esque style. The voice reaches a peak in a poignant outcry, "How will I sing us, out of this sorrow? Build a safe bridge, for the child, out of this Danger?" | All sense of directionality is temporarily lost as we are guided primarily by the changing directions of the lead vocal. Engulfed in the moving statues, lit only occasionally by strobe-like lighting bursts, this section is slightly disorienting. Especially on first viewing, the density of sound and visuals is overwhelming. |
| **(3:50 – 4:49)**  | Suddenly, the visual field turns entirely white as one's eyes adjust to the intense daylight upon leaving the dark. Björk appears in front of the viewer, now larger than life and stylized in translucent pastel purple shades. Gradually, a magical, psychedelic Icelandic landscape is revealed: mountains in the background are offset by volcanic rock on the ground, and yellow northern lights across a purple sky. | As the erratic strings fade, they give way to a long, consonant tremolo on the high strings before transitioning to luscious steady chords. Right away, the high strings are supported by thick synth pads. Now the voice is notably calmer in tone, both musically and lyrically, as the material becomes lush, "I raise a monument of love. There is a swarm of sound." | As the climax is reached, the whiteness is at first blinding, and thereafter calming light with coloration is experienced. As Björk's body reappears and she sings, the viewer is reoriented towards her. |

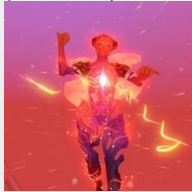| (4:50 – 5:14) | After a while, Björk's purple body disappears, and the landscape becomes more visible. Now the purples are replaced by dark greens and oranges. A black rocky sculpture like those seen earlier in the cave comprises the surface material, but with a dripping, purple "wound" vertically across the chest of the figure. Disembodied arms resembling those of the viewer's begin conducting movements above the sculpture. | The music repeats the previous phrase. The sound of wind starts to become audible, matching the heightened tangibility of the visual scene. | Musically, the strings and the synths which play in the same ranges enter from all directions, often seamless in their combination. The wind sounds also move past the listener, from one direction to another. These wind sounds are filtered in such a way that they lose their high end as it moves away. The effect is realistic. |
|---|---|---|---|
| (5:15 – 5:29) | Björk is 're-born' as she rises from the statue like the phoenix from the ashes. The object, now clearly a kind of wound on her chest, turns into a glowing source of light. The body becomes technicoloured with bright orange streamers now flow from the light in her chest. | Lyrical cue: "It will make us part of, this universe of solutions, this place of solutions, this location of solutions." | In this brief moment, the voice is solo, and experienced quite wide laterally. As Björk's body raises from the ground, it moves up continually. Once again, streamers pour out, with the viewer controlling their flow with the hands. |
| (5:30 – 6:09) | Having risen from the statue, Björk begins to walk slowly toward the viewer, as the viewer moves backward through the magical landscape. The medusa-like tentacle structure from the cave scene has returned, now framing the viewer from behind. As she walks, Björk is performing the same slow, conducting gesture with her hands. | Suddenly, the wind turns gusty as a cacophony of vocal harmony and counterpoint joins the lead voice. This point of multivocality becomes totally immersive as the visible singer walks slowly toward us. | Contrapuntal and harmonic vocal lines return, now totally consonant. They completely immerse the listener, the sensation being of warmth and comfort, rather than the fear and angst experienced in the cave. Visually, we are surrounded again by a purple tentacle-like structure, while the control over the streamers encourages dance-like motions. |
| (6:10 – 7:19) | At this stage, Björk's body has transformed its colour palette to various shades of deep red, purple, and orange. The colours resemble a sunset, the sky changing to reflect this event. She now looms larger than life, beginning to glide through the viewer through this section. | By now, the singing ceases, with high strings bending between pitches at a medium pace are clearly audible above the sound of the wind and the luscious synth pads, which sound as if they may actually be the filtered sound of synthesized voices. | In this section, Björk's body literally moves through the viewer, and in so doing prompts an urge to avoid this encounter. However, it is inevitable that she will subsume the viewer. For an instant, the viewer is encouraged to turn around and experience the scene from her perspective. |
| (7:20 – End) | As sight of the landscape dissipates, we are left with a deep purple and red haze. Björk's body, now hovering behind the viewer, slowly dissolves with the music, as the sets ends black. | As the piece draws to a close, the high strings slowly filter out, and are replaced by a sawtooth-style synthesizer playing the same line. The synthesizers and the wind gently fade out as the visuals fade into black. | Imagery and music dissolve into the sunset, at which point the viewer is rotated around Björk, who has walked through them. In the last moments she turns around to face the viewer before fading to darkness. Upon completion, when one removes the VR headset, they find themselves standing completely "backwards" of the starting position—a deliberate sense of disorientation seems to eb the objective here! |

## Immersion, Agency, and Aesthetics in Björk's VR video 'Family'

Björk belts out, "I raise a monument of love, there is a swarm of sound, around our heads, and we can hear it," as she reaches the climactic moment of the track 'Family'.[29] The lush shimmer of the Penderecki-like strings and the darting beats at this climactic point encapsulate the album's title, *Vulnicura*, namely, to be vulnerable and to be cured. This moment in the song comes across epic, a moment of transcendence when the mist lifts and the material gleams; the sonic landscape is ethereal, eloquently designed to create the sensation of a healing effect. Collaborating with Andrew Thomas Huang, Björk would experiment with digital VR technologies to produce a music video that vividly expressed her immersive experience. Huang has described how he designed the set and objects of the video:

> With a drawing and painting background, that's something I can do quite easily. It's really enriching, whereas shooting 360° video is more of a documentary-like workflow. For me, the 360° video is interesting because you are seeing the world captured as it is, untouched. Ideally with you erased.[30]

The effect of the motion-capture of Björk throughout the VR video is hyperreal, conjuring up notions of traversing magical landscapes (which were the actual landscapes used on the sets of 'Black Lake' and 'Stonemilker' shot in 360° video in Iceland). Baudrillard's theories of simulacra spring to mind when interpreting the engagement with representations of reality.[31] In a sense, hyperreality involves a simulation of reality and virtual immersion that operates more as real than the real (read: hyperreal); indisputably, Björk's VR performance creates this impression of a heightened reality. The narrative of traveling and searching is a veritable magical mystery tour, where the protagonist entices the viewer into her world by various means of identification. Compositional design, in both the imagery and music, elevates impressions of valleys, fjords, caves, open skies, and mountains. Designed by James Merry, the digital aesthetic of an artificial representation of nature is highly expressive. Björk's larger-than-life space-forms

---

[29] In chapter 10 of the Routledge research companion to popular music and gender (2017), Freya Jarman has undertaken one of the first studies into the phenomenon of belting out in popular music singing.
[30] https://www.vice.com/en_us/article/yp58dg/bjork-teases-family-virtual-reality-film-visuals
[31] See Baudrillard's essay 'The Precession of Simulacra', from *Simulacra and Simulation* (1981).

in 'Family' correspond to the viewer's own journey, where the sense of travelling in space enhances the digitalized spectacle of nature. We now turn to features linked to compositional design, in particular immersion and virtual reality, audiovisual creativity, impact of agency, and VR aesthetics These have direct correlation to the spatialities described within the VAVS model and form critical points of reference in our analysis.

### *Immersion and The Aesthetic Space*

Music contributes to a powerful sense of presence in the VR experience, with the recording furnishing an aesthetic space. There is little doubt that Björk, as artist-composer, entered into this project with an acute awareness of this and a high degree of sonic spatiality. The video's narrative unfolds as part of a game-world where actions are played out by the main character and framed by psychedelic artwork. Immersion is achieved by a lateral sense of motion that is constantly fluid – the depiction of an Icelandic landscape as a dreamworld through which Björk travels blurs the distinction between fantasy and reality. At the climactic moment (3:50), the world transforms from a dark cave to a magical purple and yellow psychedelic-tinged surround, with Björk's digital body moulding into the same palate as her surroundings. If and when looking upwards, the viewer perceives what resembles a shimmer of yellow 'northern lights', spanning the purple background, signalling both a shift in the subjective perspective and the magical scenes of Iceland. Importantly, the temporal displacement of the sound image in the overall experience functions to generate various impressions of an active environment.

Immersion, in this instance, is mediated by Björk's agency and predicated upon a host of intricate details. In terms of the audiovisual space, Björk's gestures literally reach out to the viewer (for instance, from 5:30 onwards), beckoning them to make contact with her virtual hands. Useful here is William Moylan's notion of 'lateral imaging' that describes the placement of sonic objects in the sound stage, as well their perceived size and width (Moylan 2012, p. 176). In 'Family' Björk's larger-than-life presence at particular moments (e.g., at 3:06, 3:50, 5:30, 6:10) is contingent on expansions in the apparent size of her voice(s) in the 3D mix. The viewer is drawn into the aesthetic space through their own movement in the lateral plane; the result is that of feeling lost in space. This phenomenon is heightened prior to the 'cave climax' (3:06) when the absence of the singer in the forefront creates a confusing laterality causing the viewer to search their surroundings, and

also towards the end of the video (~6:10) as Björk's digitized body glides literally through the viewer. With subtle movement and interaction, the viewer increases not only their propensity for immersion through the agential space, but also their perception of the surrounding lateral spatialities created both by the artist and viewer.

From this we want to suggest that part of what constitutes aesthetic space lies in the perceived proxemic relationship between performer and viewer, since the viewer is in constant interpretive negotiation between their own subjectivity and that of the artist. There are specific ways in which the construction of the sonic space creates new proxemic relationships in the aesthetic space. For instance, the 3D mix in virtual reality allows for a placement of reverb and delay surrounding the listener, creating proxemics which are simultaneously perceived as intimate while retaining vast and lush reverb and delay profiles. This is certainly the case in the aforementioned 'cave climax' (3:06), where the reverb on Björk's voice is panned opposite to her position in the 3D mix. Later in the piece, there are moments when the counterpoint vocal lines are as loud or louder than the main vocal line which tracks the singer's digital body. These vocal parts are afforded various spatialisation profiles, from the feelings of whisper in the ear to distant repetitions of the lead voice's lyric and melody. In her analysis of the Björk song 'Vespertine', Dibben states that "the lyrics are simultaneously intimate and self-revealing such that they accomplish a striking alignment of the sensual with the spiritual" (2007, p. 176). Similarly, the lyrics in 'Family' suggest such tendencies, not least in the intimacy of Björk's vocal sound and its skilful panning in the 3D mix. As in much of her work, Björk turns to structures of intimacy through the minute details of her recorded voice, and in the 3D mix this is embodied in the background voices which become inner thoughts, whereby the cacophony of vocal textures has a sense of reflective and emotional inner dialogue.

Immersive sensations, such as those described above, establish a mutual space for performer and listener, the purpose being to communicate a sense of musical passage. Impressions of changing spatiality thus establish a holistic sensation of an environment that can pull us in any direction. In this case, agential space results from contrasting interfaces of colour, imagery, sound, bodily gestures, and above all, the three-dimensional surrounding of the viewer. In sum, the virtual reality experience of the narrative of 'Family' is conditional on the merging of physical

and cerebral interaction that works through a pop art aesthetic,[32] dependent on constant transformation and innovation.

## *Audiovisual Creativity: Voice and Visual Space*

In the video Björk traverses a barren Icelandic landscape: imagery changes constantly align with musical events, the moment of transcendence occurring at 3:50, where the performer arrives at a 'gateway of enlightenment' suddenly drenched in a swarm of sounds. Structurally, the song's sections are relegated to visual happenings. Carol Vernallis (2019) has observed that song-sections unfold according to narrativity, and within the frames of each section of 'Family' Björk's agency can be assessed according to creativity. Overtly, she constructs her persona around a personal narrative that arguably possesses a high degree of authenticity.[33] On the function of the musical persona, Phil Auslander has stressed that an artist's appearance concerns "the visual dimensions of self-presentation, while manner has to do with the behavioural dimension" (2019, p. 96). With Björk, her persona is reinforced by a performance that is instantly identifiable as genre-specific in terms of trademark; her mannerisms and visual traits affirm an expression that is familiar to any of her fans, which directly mediates her very own digital signature. Ample opportunity to explore the imagination of Icelandic landscape, in tandem with her own quest for clarity, is on offer for the viewer.

Sonically, the treatment of the voice reinforces Björk's visually presented intimacy, both in relation to the viewer and to her connectedness with the imagery of Iceland on display. Although the motion-capture images of Björk's body are at times distorted, garishly coloured, and heavily stylised, the voice often retains a relatively dry and intimate quality. Close inspection in the sonic space, for example from around 6:00 until the end, discloses her voice dubbed and mixed in with luscious strings and synthesizer textures, coming across as heavily processed 'choirs'. In one sense, these choral textures serve as a connection point between the sonic and visual spaces, creating an aesthetic bridge from the sounds of strings to a surrealist landscape, all depicted by the voice. This also serves to confound

---

[32] By pop art aesthetic we are referring to both the surreal, profound and banal (Hawkins 1997) that stems from pop art's beginning in the mid to late 1950s where artists, such as Roy Lichtenstein, Andy Warhol, Jasper Jones, Tom Wesselman and others, derived their inspiration from subject matter found in everyday popular culture.
[33] For theories on personal narrative see Hawkins & Richardson 2007, Hawkins 2020.

the processes of source-bonding, since the sonic boundaries between choral overdubbing and string instruments are purposefully blurred.

Perhaps the most profound feature in the video is the sculpting of a space that is sonically expansive through technologies of spatialization. Tensions in the sonic materiality achieve different senses of space, where a wide array of sounds are constantly mobile; they are charged technologically through the details of production. Compared to the majority of her music videos, the VR technology employed in 'Family' arguably turns Björk into a 'virtual star', with a set of immersive qualities denoting a high degree of exceptionality. While the personal narrative might seem overt, there is an impression of a staged fictitious persona at work due to the 3D projections and communicative options open for the viewer to enter the set.

### *Impact of Agency: authorial intent*

Common to Björk's *oeuvre* is a sense of full control of performance and production. In this sense, her 'authorial intent' within a transpersonal space (Hawkins 2002, pp. 15-16) can be assimilated against the authentic representations of her own attitude to performance. How then does the VR music video contribute to the relationship between viewer and singer? And what strategies are negotiated to facilitate a virtual sense of staging on the part of the listener as much as the performer? The impact of her performance results from the practice of signifying 'reality' in terms of the 360° video. Aspects of visual spatialization in 'Family' merge into a sonic soundworld where Björk's voice is foregrounded as intimate (Dibben 2012; Kraugerud 2020). Intimacy becomes a primary sonic device for drawing attention to the narrative and lyrical meaning. By creating a life-like persona in 'Family,' a sense of hyperembodiment[34] via the VR sensation is like being physically in touch with the artist. The multitude of positions offered up in the video are striking markers of agency, and one way to comprehend this is through multimodality where the composite of the performance is a result of different expressive modes. Burns has theorized this through 'expressive channels' or 'domains' that can be summed up as 'word-music-image' (2019, p. 184).

---

[34] Hyperembodiment is theorized by Stan Hawkins in an analysis of Rihanna's music video 'Umbrella' from 2007, where it is argued that an obsession with the look is conditional on technologies of musical production (See Hawkins 2013, p. 481). Also see Kai Arne Hansen's analysis of Beyoncé's sonic staging of the gendered body as a means to foregrounding hyperembodiment as a mechanism of digital fetishization (Hansen 2017).

Ultimately, Björk's embodied gestures guide the audiovisual aesthetics, brought into focus by the processes of production. As such, her corporeality is supported by the composite of word-music-image, which discloses an array of strategies.

One might say that the sense of journey in the VR experience entails a trajectory of author-induced imagery, inspired by the finely detailed audiovisual aesthetics. Björk's agency, a prime constituent and determinant of the compositional design, is aided by techniques of temporal regulation made realistic by close-up shots of her gestures and the merging of her with the viewer at specific points (for instance, at 6:10). Regulated integration with the viewer accomplishes a strong sense of identification, facilitating the pleasurable aspect of spectatorship. From 5:40 onward this is intensified as Björk begins a repetitive hand motion that resembles a kind of slow, ethereal conducting, or perhaps a sewing motion. This hand gesture can be performed and transferred to the viewer throughout the video.

Regulated by a button on the VR controller, in our case a pair of controllers for our Oculus Rift headset, the viewer's hands move in the same gesture and ensure interaction. Emphasis falls on Björk's agency as a performer is mediated through technologies of spatialisation as much as on the agency of the viewer. By the end of the recording, the impact of immersion is at its height as the viewer removes the headset; if they have 'followed' Björk's digitised character throughout, they discover they are facing the wrong direction—turned around 180 degrees from where they faced at the beginning. In such an instance of disengagement with the media, the viewer likely realises their own interaction with the song.

Worth emphasizing is the technology of viewing the VR video itself, which comes across with its own set of agential limitations. Jacquelyn Ford Morie has referred to the "bifurcated self", wherein "the act of emplacing one's body within the immersive environment signifies a shift into the dualistic existence in two simultaneous bodies" (2007, pp. 127-128). This is certainly the case in viewing the video on an Oculus headset and controlling one's digital arms with motion-sensing controllers—one is both the embodied character in a VR video and a person viewing the video, aware of the technological distance between the two selves, but feeling nonetheless connected to both.

### VR Aesthetics

In probing further at VR aesthetics, we would suggest that the link between reality, hyperreality, and virtual reality is made tangible by optic arousal and strategies of representation. A wide palette of colours veer towards the florescent and garish at

times, enhanced by the use of lighting, which continually helps paint the environment, hence intensifying the staging of Björk's performance. Within an immersive VR environment, colours and lighting contribute to the perception of subjectivity. More specifically, nuances in the technical and stylistic coding of the compositional create visual impressions of light and shade. Correlating with timbre, texture and dynamics are finely regulated hues: blue merges with pink, orange turns into green, and so on. In addition, changes to sonic spatialization emphasize nuances of colour, signifying attitudinal and emotional content in the subject matter.

In assessing the full effect of the aesthetics of 'Family', we return to the question of the hyperreal and the digital simulation of reality both in the music and in the visuals. Narratively, there is a sense that Björk escapes the real world by journeying into the hyperreal one: the entire edifice of the digital production is reliant on her calling into question her own reality. It is within the confines of such spatiality that Björk's mission becomes most compelling. After all, the song is about her mourning the loss of her relationship to the US artist Matthew Barney in heart-wrenching lines, such as "Is there a place, where I can pay respects, for the death of my family." Ever so poignantly, her video stands as an unrivalled testament for expressing such pathos.

## Conclusion

In concluding, we would like to return to the issue of staging and immersion in VR music videos and the cognizing of spaces, shapes, and designs. Dibben's in-depth study (2013) of Björk's approach to digitalization in *Biophilia* (2011) revealed the profound changes and effects a mobile app format had on the way people experience music. Introducing multimodality and interactivity to the experience of recorded music, the emergence of new aesthetic implications for the visualization and immersive modes of listening warrants attention. As Dibben insists, mobile music apps have formed a medium that offers interactive functions that lead to creative versions of Björk's songs. In opting for touch screens, Björk aimed for a new creative experience that combined technology, interactivity and nature. A distinctive feature of this was the integration of concept and aesthetic, which, first, encouraged a visualization of music in a way that encouraged "attentive listening to and playing with musical structures and processes," second, offered "a multimodal experience by virtue of touchscreen interactivity," and, third, presented "a curated experience of a coherent artistic vision" that was the result of

collaborative work (Dibben 2013, p. 688). A major consequence of the audiovisual relationships emerging from touchscreens was not only a renewal of modes of listening, but also a spontaneously embodied mode of engagement.

The powerful creative vision of collaborative music products by Björk would some years later be extended into the VR music format, which we have addressed in this article, where a host of new possibilities for comprehending compositional design are evidenced in our discussion of this format. As our model of virtual audiovisual space (VAVS) exemplifies, the listener's position and role merges with the act of staging a performance in an innovative format that extends the app album designed for mobile digital devices. One could argue that the audiovisual analysis of surround and 3D sound gets us to ponder over the developments and intricacies in music production on a broader scale. This is because the staging of sonic and visual devices in VR music videos allow for a greater sense of interaction between artist and fan, and in this dialogic space intertextual pathways are (re)invented and constructed. The conception here is that the new digital medium of music VR experience incorporates immersive functions that align pop music more to computer games. As with the user of games, the user of VR music videos has a wide scope to interact and perform along with the artist.

Ultimately, Björk's 'Family' VR video addresses features found in the artist's earlier work as we revisit her relationship between nature and technology through an acute practical engagement. Such fascinating technological innovations also give good cause for re-examining the normative structures that define pop dramaturgy, providing an opportunity to probe at the advances in technology and ponder over the future of new audiovisual aesthetics. Moreover, the multimodal aspect of 'Family' illustrates a coherent vision of compositional and performance design that has significant implications for understanding pop aesthetics and the phenomenon of music making. It is our hope that future studies will engage with the particularly nuanced phenomena of VR immersion, as a new generation of music video production continues to affect human development, agency, and creative expression.

# References:

Altman, R. 1992. Sound Space. In: Altman, R. (ed.), Sound Theory / Sound Practice. London: Routledge.

Auslander, P. (2019). Framing Personae in Music Videos. In L. A. Burns & S. Hawkins (Eds.), The Bloomsbury Handbook of Popular Music Video Analysis (pp. 91-109). New York, NY: Bloomsbury Academic.

Baudrillard, J. (1994). Simulacra and simulation (S.F. Glaser, Trans.). Ann Arbor: University of Michigan Press. (Original work published 1981).

Bolter, J. D., & Grusin, R. (2000). Remediation: Understanding new media. Cambridge: MIT Press.

Bresler, Z. (2021). Immersed in Pop: 3D Music, Subject Positioning, and Compositional Design in The Weeknd's 'Blinding Lights' in Dolby Atmos. Journal of Popular Music Studies, 33(3).

Burns, L. A. (2018). Interpreting Transmedia and Multimodal Narratives: Steven Wilson's "The Raven That Refused to Sing". In C. Scotto, K. M. Smith, & J. Brackett (Eds.), The Routledge Companion to Popular Music Analysis: Expanding Approaches (pp. 95-113). New York and London: Routledge.

Burns, L. A. (2019). Dynamic Multimodality in Extreme Metal Performance Video: Dark Tranquillity's 'Uniformity', Directed by Patric Ullaeus. In L. A. Burns & S. Hawkins (Eds.), The Bloomsbury Handbook of Popular Music Video Analysis (pp. 183-200). New York, NY: Bloomsbury Academic.

Burns, L. A., & Hawkins, S. (2019). The Bloomsbury Handbook of Popular Music Video Analysis. New York, NY: Bloomsbury Publishing USA.

Camilleri, L. (2010). Shaping sounds, shaping spaces. Popular Music, 29(2), 199-211.

Chion, M. (1994). Audio-vision: Sound on Screen (C. Gorbman, Trans.). New York: Columbia University Press.

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

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

Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. New York, NY: Harper Perennial.

DeNora, T. (2000). Music in everyday life. Cambridge: Cambridge University Press.

Dibben, N. (2007). Subjectivity and the Construction of Emotion in the Music of Björk. Music Analysis, 25(1), 171-197.

Dibben, N. (2012). The Intimate Singing Voice: Auditory Spatial Perception and Emotion in Pop Recordings. In D. Zakharine & N. Meise (Eds.), Electrified Voices Medial, Socio-Historical and Cultural Aspects of Voice Transfer (pp., 107-122). Göttingen, DE: V&R unipress GmbH.

Dibben, N. (2013). Visualizing the App Album with Björk's Biophilia. In C. Vernallis, J. Richardson, & A. Herzog (Eds.), The Oxford Handbook of Sound and Image in Digital Media (pp. 682-704). New York: Oxford University Press.

Douglas, Y., & Hargadon, A. (2000). The pleasure principle: immersion,

engagement, flow. Proceedings of the 11th ACM on Hypertext and Hypermedia. San Antonio, TX.

Hansen, K. A. (2017). Fashioning Pop Personae: Gender, Personal Narrativity, and Converging Media in 21st Century Pop Music. (Ph.D). University of Oslo, Norway.

Hawkins, S. (1997). The Pet Shop Boys: Musicology, masculinity and banality. In S. Whiteley (Ed.), Sexing the Groove. London: Routledge.

Hawkins, S. (2002). Settling the Pop Score: Pop texts and identity politics. Burlington, VT: Ashgate.

Hawkins, S. (2013). Aesthetics and Hyperembodiment in Pop Videos: Rihanna's 'Umbrella'. In J. Richardson, C. Gorbman & C. Vernalis (Eds.), The Oxford Handbook of New Audiovisual Aesthetics (pp. 466-482). Oxford: Oxford University Press.

Hawkins, S. (2020). Personas in rock: "We Will, We Will Rock You." In A. Moore and P. Carr (Eds.), The Bloomsbury Handbook of Rock. New York, NY: Bloomsbury Publishing USA (forthcoming).

Hawkins, S. & Richardson, J. (2007) Remodeling Britney Spears: Matters of Intoxication and Mediation. Popular Music and Society, 30(5), 605–629.

Herbert, R. (2011). Everyday Music Listening: Absorption, Dissociation and Trancing. Surrey, UK: Ashgate.

Høffding, S. (2019). A phenomenology of musical absorption. London: Springer.

Idrovo, R. & Pauletto, S. 2019. Immersive Point-of-Audition: Alfonso Cuarón's Three-Dimensional Sound Design Approach. Music, Sound, and the Moving Image, 13, 31-58.

Jamieson, G. A. (2005). The Modified Tellegen Absorption Scale: A Clearer Window on the Structure and Meaning of Absorption. Australian Journal of Clinical and Experimental Hypnosis, 33(2), 119-139.

Jarman, F. (2017). High Notes, High Drama: Musical climaxes and gender politics in tenor heroes and Broadway women. In S. Hawkins (Ed.), The Routledge Research Companion to Popular Music and Gender (pp. 137-151. New York: Routledge.

Jenkins, H. (2006). Convergence culture: where old and new media collide. New York: New York University Press.

Kelly, J. (2019). The Palimpsestic Pop Music Video. In L. A. Burns & S. Hawkins (Eds.), The Bloomsbury Handbook of Popular Music Video Analysis (pp. 219-233). New York, NY: Bloomsbury Academic.

Korsgaard, M. B. (2013). Music Video Transformed. In J. Richardson, C. Gorbman & C. Vernalis (Eds.), The Oxford Handbook of New Audiovisual Aesthetics (pp. 501-521). Oxford: Oxford University Press.

Korsgaard, M. B. (2017). Music Video After MTV: Audiovisual Studies, New Media, and Popular Music. New York and London: Routledge.

Kramer, L. (2011). Interpreting Music. Berkeley: University of California Press.

Kraugerud, E. (2020). Come Closer: Acousmatic Intimacy in Popular Music Sound (PhD thesis, University of Oslo).

Krueger, J. (2009). Enacting musical experience. Journal of Consciousness Studies, 16(2-3), 98-123.

Liljedahl, A. A. (2019). Musical Pathfinding; or How to Listen to Interactive Music Video. Music, Sound, and the Moving Image, 13(2), 165–85.

Morie, J. F. (2007). Performing in (virtual) spaces: Embodiment and being in virtual environments. International Journal of Performance Arts and Digital Media, 3, 123-138. doi:10.1386/padm.3.2-3.123_1

Moylan, W. (2012). Considering space in recorded music. In S. Frith & S. Zagorski-Thomas (Eds.), The Art of Record Production: An Introductory Reader for a New Academic Field (pp. 163-188). Surrey: Ashgate.

Murray, J. H. (2016). Hamlet on the Holodeck: The Future of Narrative in Cyberspace (2 ed.). New York: The Free Press.

Rambarran, S. (2021). Virtual Music: Sound, Music, and Image in the Digital Era. New York: Bloomsbury Academic.

Richardson, J., Gorbman, C., & Vernallis, C. (Eds.). (2013). The Oxford Handbook of New Audiovisual Aesthetics. Oxford University Press.

Rogers, H. (2011). The Unification of the Senses: Intermediality in Video Art-Music. Journal of the Royal Musical Association, 136(2), 399-428.

Rogers, H. (2014). Spatial Reconfiguration in Interactive Video Art. In K. Collins, B. Kapralos, & H. Tessler (Eds.), The Oxford Handbook of Interactive Audio (Online) (1 ed.). Oxford: Oxford University Press.

Sanders, T., & Cairns, P. (2010). Time perception, immersion and music in videogames. Paper presented at the Proceedings of the HCI International 2009, San Diego, CA.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of The Royal Society B, 364, 3549-3557.

Smalley, D. (2007). Space-form and the acousmatic image. Organised Sound, 12(1), 35-58.

Smalley, D. (1997). Spectromorphology: explaining sound-shapes. Organised Sound, 2(2), 107-126.

Vernallis, C. (2019). Writing about music video. In L. Patti, ed. Writing About Screen Media, New York and London: Routledge.

Vernallis, C. (2013). Unruly Media: YouTube, Music Video, and the New Digital Cinema. Oxford: Oxford University Press.

Wishart, T. (1996). On sonic art (2nd ed.). (First edition 1985). Amsterdam: Routledge.

Wolf, W. (2015). Literature and Music: Theory. In G. Rippl (Ed.), Handbook of intermediality: Literature–image–sound–music (Vol. 1). Berlin: Walter de Gruyter GmbH & Co KG.

**Audiovisual Reference:**

Björk. (2019). Vulnicura VR [VR Album]. UK: One Little Indian, Analog Studios. Available on Steam: https://store.steampowered.com/app/1095710/Bjrk_Vulnicura_Virtual_Reality_Album/ (Downloaded Oct. 2019).