Tracing Embodied Narrative in VR experiences

Szilvia Ruszev*

University of Southern California (US)

Abstract

In this article, I will consider the specificity of narrative in VR from the perspective of its embodied, spatial, and participatory nature. For this purpose, I will look at storytelling in VR from a neurofilmological perspective that accounts for both the cognitive and phenomenological and conceptualizes the viewer as an organism, applying an integrative view. I will probe several ideas from cognitive sciences and phenomenological cinema and media studies in order to support the above understanding of narrative in VR. I will apply these theories and ideas to specific VR case studies: *Book of Distance* (Randall Okita, 2019), *Heterotopias* (Noa Kaplan and Szilvia Ruszev, 2018), and *Carne y Arena* (Alejandro Iñárritu, 2017). Finally, this article will discuss the terms "ambient storytelling" (Stein and Fisher 2013), "somatic montage" (Waite 2016), and "embodied narrative" (a term used in cognitive sciences) and their relevance concerning narrative in VR.

Keywords: Embodied narrative; Somatic montage; Ambient storytelling; Event segmentation; Neurofilmology.

Szilvia Ruszev: University of Southern California (US)
✉ ruszev@usc.edu
Szilvia Ruszev is a film editor, media artist and scholar working across different media formats. Her broader research interest focuses on sensuous knowledge, montage theories and politics of post cinema. Her own artistic work relates to very personal moments, certain states of emotional solitude in relation to the Other, both in its particular and abstract notion. As editor, she collaborated with internationally acclaimed directors such as Peter Greenaway, Anders Østergaard, and János Szász. Her award-winning work has been part of numerous international film festivals and exhibitions such as Karlovy Vary IFF, TIFF Toronto, Berlin IFF, Siggraph and Codame. Between 2010-2016, she was a faculty member of the Editing Department at the Film University Babelsberg Konrad Wolf. Currently, she is an Annenberg Fellow, pursuing a Ph. D. degree in Media Arts + Practice in the School of Cinematic Arts at the University of Southern California.

∗　✉ ruszev@usc.edu

1   Introduction

The medium of Virtual Reality (VR) has been swiftly expanding in the field of entertainment as the technology behind it becomes more affordable and comfortable to wear. This medium has been sustained by the great expectation of total immersion in an expanded reality, articulated by VR guru Jaron Lanier in an electrifying way:

> VR is one of the scientific, philosophical, and technological frontiers of our era. It is a means for creating complete illusions that you're in a different place, perhaps a fantastical, alien environment, perhaps with a body that is far from human. And yet, it's also the farthest-reaching apparatus for researching what a human being is in terms of cognition and perception (Lanier 2017: 1).

Yet immersion and presence, the desires most often expected from VR, are not new desires. Humans have created illusory visual spaces since the beginnings of human culture (Grau 1999). Physical spaces have been visually altered to create enclosed spaces of ritual, private or public political action, as Oliver Grau has pointed out in his article "Into the Belly of the Image." These experiences, he argues, "nestle into viewers' senses" and create a totality of immersion (Grau 1999: 365). The intensity of this artificially created sensual sheath aids the fundamental human activity of storytelling through which we make sense of the world.

Investigating VR as a narrative medium, we must consider its specific affordances that shape how narrative can be conceptualized. VR has been discussed as an embodied, interactive, and user-oriented medium (Blach 2008). In other words, a VR experience is centered around the viewer and their body and it unfolds through the interaction with or exploration of the narrative by the viewer, who physically shares the time and space of the narrative (Aylett and Louchard 2003: 3).

In this article, I will consider the specificity of narrative in VR from the perspective of its embodied, spatial, and participatory nature. For this purpose, I will look at storytelling in VR from a multidisciplinary perspective that accounts for both the cognitive (*viewer-as-mind*) and phenomenological (*viewer-as-body*). "Neurofilmology," as described by Adriano D'Aloia and Ruggero Eugeni (2014), conceptualizes the viewer as an organism, applying an integrative view.

I am proposing an understanding of VR as a multisensory, embodied, and spatial medium of virtuality. The viewer-organism becomes an explorer in a spatial narrative experience in which they actively encounter compositional elements and charter a unique, multisensory, and embodied constellation. The body, through the senses of kinesthesia and proprioception, becomes a sensitive storytelling device. Therefore, narrative in VR must be viewed as an embodied, spatial and participatory process that unfolds through the viewer's exploration (Aylett and Louchard 2003: 1). For the purposes of this article, I will probe several ideas from cognitive sciences and phenomenological cinema and media studies in order to support the above understanding of narrative in VR.

On the cognitive side, I will look at Tim J. Smith's (2012) "attentional theory of cinematic continuity," through which he discusses how visual attention plays a role in the perception of continuity across cuts in movies. I argue that this theory can be applied equally well to VR as visual attention plays a crucial role in guiding the viewer's exploration toward a continuous narrative experience. Further, I will discuss the so-called "event segmentation theory" (Zacks, Speer et al. 2007), which I argue is at the basis of "storifying" an experience. The theory of "embodied simulation" discusses how "part of the neural resources that are normally employed to interact with the world around us, shaping our relationships and relations, are reused for perception and imagination" (Gallese and Guerra 2015: 1). This theory of embodied simulation has been applied in the field of cinema studies as a way to make sense of the perceptual processes used while watching movies and can be expanded to understand VR. Alongside mental simulation processes, body-related senses such as kinesthesia and proprioception play a significant role in exploring and shaping the VR narrative.

On the phenomenological side, I will look at the embodied character of formal cinematic elements such as color, movement, camera angle, mise-en-scène, sound, and editing and how they impact the viewer on an affective, bodily level. Scholars such as Vivian Sobchack (2010) and Tarja Laine (2018) have extensively written about the reciprocal relationship between the material character of the film and the viewer. This kind of

affective reciprocity can be applied in the case of VR and plays a significant role in the active "storification" process fulfilled by the viewer. I will apply these theories and ideas to specific VR case studies. I will focus on three VR experiences: *Book of Distance* (Randall Okita, 2019), *Heterotopias* (Noa Kaplan and Szilvia Ruszev, 2018), and *Carne y Arena* (Alejandro Iñárritu, 2017).

For its narrative design, the VR experience *Book of Distance* explores the spatial and aesthetic character of the theater, inviting the viewer not only to observe but to interact with some objects. It is a documentary-infused experience in which the author (appearing himself as an avatar using his own voice) guides us through his family's story of immigrating from Japan to Canada and living through harrowing times during the Second World War. The VR experience uses existing artifacts such as family photos, newspaper excerpts, and models of objects that bear importance to the family's personal history. The viewer is invited into a virtual space of memories by the author-narrator, who is also present. The viewer is prompted to discover and interact with objects that are significant in the narrative. They can grab the digital facsimile of family photos, documents, and old newspapers and look at them closely or interact with them in other ways. These gestures function as triggers in advancing the narrative on the one hand and adding an emotional depth to a non-fictional narrative created in VR on the other.

Heterotopias is a VR essay based on Michel Foucault's influential lecture, *Des espaces autres*. In Foucault's (1986) words, "the heterotopia is capable of juxtaposing in a single real place several spaces, several sites that are in themselves incompatible" (23). The primary goal of this VR essay is to establish a sensory version of Foucault's heterotopology, which he defines as "the study, analysis, description, and 'reading'" (23) of places that simultaneously reflect and invert the rules of social engagement. The experience uses developing eye-tracking technology to transform users' blinks into cinematic cuts. With every blink, the virtual space alters, producing variable configurations of stereoscopic 360-degree footage and computer-generated models. To complement the ungrounded transformation of space, the series includes spatialized audio and custom-made furniture. The experience is supplemented by spatial audio of an ambient soundscape, including a whispering sound collage based on the English translation of Foucault's original lecture (Kaplan and Ruszev 2018). *Heterotopias* doesn't employ storytelling in its traditional understanding, as it lacks both characters and action. A feeling of a narrative arises from the succession of the spaces – the well, the garden, the movie theater, the cemetery, the mausoleum, and the mirror – and their symbolic and embodied metaphorical meaning.

Alejandro Iñárritu's *Carne y Arena* is an immersive installation based on the accounts of refugees who have been trying to cross the between the United States and Mexico. The installation consists of three stages. In the first stage, the viewer is asked to enter barefoot a metal-walled and freezing cold room, evoking a detention center. In the second stage, the viewer experiences the border crossing situation in VR. And in the third stage, the viewer is confronted with the portraits and stories of the real people on which the experience and characters are based.

Finally, this article will discuss the terms "ambient storytelling" (Stein and Fisher 2013), "somatic montage" (Waite 2016), and "embodied narrative" (a term used in cognitive sciences) to discuss their relevance concerning narrative in VR.

2   Defining VR and the viewer-explorer

The definition of VR itself is not necessarily standardized in the literature. There is a more pronounced distinction between the so-called 360-degree videos or cinematic VR and "proper" VR experiences, but for the sake of this contribution, I will be referring to VR in its traditional sense.[1] Various definitions agree that VR is an immersive, multisensory, interactive, computer-generated, and three-dimensional viewer-centered ex-

---

1.   Cinematic VR has only 3DoF (three degrees of freedom) and the viewer cannot move or interact with the virtual space, instead they merely look around in a 360-degree sphere and, if available, have some basic interaction with the video (stop, play, "jump" to a next scene). In its traditional sense, VR has 6DoF (six degrees of freedom), which means that the viewer can move around in the virtual space so that objects in the space will appear closer or further away. Moreover, the user has the possibility to interact with elements of the virtual environment depending on how the experience has been designed and what input devices can be used beyond the VR headset.

perience (Mazuryk and Gervautz 1996). In her book *Narrative as Virtual Reality*, Mary-Laure Ryan lists a comprehensive set of characteristics of the medium.

> 1. You enter (active embodiment)... 2. into a picture (spatiality of the display)... 3. that represents a complete environment (sensory diversity). 4. Though the world of the picture is the product of a digital code, you cannot see the computer (transparency of the medium). 5. You can manipulate the objects of the virtual world and interact with its inhabitants just as you would in the real world (dream of a natural language). 6. You become a character in the virtual world (alternative embodiment and role-playing). 7. Out of your interaction with the virtual world arises a story (simulation as narrative). 8. Enacting this plot is a relaxing and pleasurable activity (VR as a form of art). (Ryan 2001: 51-52)

VR as a medium and technology finds its application in various fields such as education and health care, alongside its territorial gain in entertainment formats such as video games and other narrative genres. In this article, I am focusing on a particular form of VR that can be found on the spectrum between cinematic VR and video gaming. Characteristic of this form is the creative use of the possibilities given by game engines. Yet, at the same time, this form strives to create a mode of second-person narration that moves this form's aesthetic and narrative choices closer to cinema.

Cinema and gaming reflect the role of the viewer in their specific terminology. Between a spectator of a movie and a computer game user, there is a whole range of possible levels of interaction and agency. For the kind of VR experience relevant here, the term "viewer-explorer" seems to be the best suited. The word exploration implies a spatial activity, "to travel over (new territory) for adventure or discovery," as the Merriam-Webster Dictionary defines it.[2] The term exploration also contains a mixture of goal-oriented pursuit and serendipitous stroll.

VR as a medium has its kinship with four other media – cinema, theater, video installation art, and computer gaming. Each of these fields has its established theoretical approaches concerning aesthetics, narrative design, and the role of the viewer. If we have to emphasize the most prominent characteristic for each medium to which VR has the closest proximity, then for cinema, it is narrative; for theater, it is presence; for video installation, it is spatiality; and for gaming, it is interactivity. In this sense, VR can be assessed as an emerging medium where elements of film, theater, video installation art, and computer gaming converge in a new medium. VR "feels" close to cinema by "inheriting" its audiovisual grandiosity and yet stays uncanny by combining this grandiosity with a very different inhabitation of space – materially as well as temporally – by placing the viewer-explorer in the center of the experience. The diegetic time and space of VR converge with the real-time experience of the viewer-explorer. Looking at VR from the vantage point of theater, especially immersive theater, brings in a participatory aspect of the "spectActor" (Boal 1999). The spectActor, in this case, is physically present and immersed in a real-time narrative that they can shape. Video installation as a form of expanded cinema and VR has similarities in how the configuration of physical space and moving image facilitates an active, explorative relationship between the viewer and the experience. Finally, computer gaming and VR cover a substantial shared territory in interactivity and dynamic systems.

3   Embodiment and VR

The notion of embodiment plays a significant role in defining narrative in VR. I argue that the characteristic way in which the body of the viewer-explorer encounters a VR experience shapes the narrative in a participatory process. The term embodiment has been used differently in various disciplines. In its philosophical perspective, embodiment refers to a general understanding of how one understands and defines oneself (Blanke and Metzinger 2009). Cognitive sciences and psychology are generally concerned with the relationship between cognition and the agent's body. Although depending on the specific field, this relationship is viewed quite differently (Graziano and Botvinick 2002; Chemero 2009).

The term "sense of embodiment" (Kilteni, Groten et al. 2012) has been used in the context of VR to distinguish it from the general notion of embodiment. It consists of the components of the sense of self-location, the sense

---

of agency, and the sense of body ownership. In the case of VR, the sense of owning, controlling, and being inside a biological body is complicated by the circumstances that it is also immersed in a virtual environment and extended with a virtual avatar.

Conceptualizing the body of the viewer-explorer and their sense of embodiment is critical in expanding the understanding of narrative in VR. In VR, the viewer-explorer's body is immersed in a perceptally hybrid environment. Stimuli are coming from both the computer-generated virtual environment and the physical world. Ideally, these stimuli are congruent. The blending of these two kinds of environments varies depending on the specific VR experience. For example, in Alejandro Iñárritu's VR experience *Carne y Arena*, the viewer-explorer is barefoot and walking on a sandy surface, which enhances the perceptual authenticity of the virtual environment of being situated in a desert. In the VR experience *Heterotopias*, the viewer-explorer's main bodily engagement with the experience is achieved with the help of a hanging chair. Through its suspension and relatively large size (a circular form with an approximate five-foot diameter), the viewer-explorer can change their posture and desired movement in the chair – they can choose to quietly sit with feet on the floor or lay down and swing around. The VR experience uses the physical affordances of the hanging chair to connect it to the VR spaces or specific objects within the given space. In the first scene, the viewer-explorer finds themselves suspended in a dark well, with a bucket dangling on a rope within the enclosed space. At the end of the first scene, the viewer-explorer gets pulled up from the dark well into a light-filled garden. An oversized birdcage hangs from the top and slowly swings over the head of the viewer-explorer. In the last space, several mirrors are suspended freely in the space of a mausoleum. They swing slowly, deliberately distorting and fragmenting the space.

In VR, the involvement of the body is determined by a perceptally hybrid environment that enwraps the viewer-explorer with both physically existing stimuli (such as the chair they sit on or the floor they are standing on) and virtual stimuli. Visual and auditory elements deliver the primary source of stimuli, which is complemented by haptic or olfactory stimuli depending on the technology involved. Notably, the body's involvement in VR activates kinesthetic (perception of one's own body parts and their movement) and proprioceptive aspects (awareness of the spatial orientation and presence of one's own body). Kinesthesia as a term – from Greek *kinein* "to set in motion; to move" and *aisthesis* "perception" – was first coined by psychologist and neurologist Henry Charlton Bastian in the late nineteenth century as the "Sense of Movement" (Bastian 1880).[3] Kinesthesia has been referred to as the "sixth sense," defined as "the perception of weight, effort and resistance, movement and position [...] involving the sensibilities of muscles, tendons, joints, and skin" (Boring 2019: 525). While kinesthesia is introspective, proprioception is outward-oriented and includes a sense of balance and orientation. Understanding the mechanisms of these body-space-body-related senses is key in developing an embodiment theory in the context of VR. Since the body of the viewer-explorer is at the epicenter of the experience, each formal, aesthetic, and story element of the experience is "evaluated" by the viewer-explorer in its relation to their own body. Layers of the virtual, physical, and biological overlap and complicate the experience by unsettling these same categories. The boundaries between one's own body and the environment, the body of the other, or one's own virtual avatar can become blurry in VR. Sita Popat argues that "bodies within these contexts may be experienced simultaneously as absent and present, together and separate" (Popat 2016: 360). In that sense, the body of the viewer-explorer oscillates between the states of being an author and a character in the VR experience. The actions in the VR experience are shaped by the viewer-explorer and, at the same time, the viewer-explorer is affected by these actions.

In *Carne y Arena*, the viewer-explorer encounters the precarity and violence of an illegal border crossing situation. The gradual crescendo of the stimuli surrounding the viewer-explorer peaks when the border patrol agents beleaguer a group of people trying to cross the border. The viewer-explorer is surrounded by agents shouting, dogs barking, and helicopters whirring above them. It is a physically unbearable situation not only because of the actions one sees, but the audiovisual effects of loud sounds and glaring lights. As a viewer-explorer, I cannot exclude my body from the stressful situation, but I can choose my physical position within it and either stay close or remove myself and observe.

The Book of Distance goes a step further by inviting the viewer-explorer to interact with some of the objects within the VR experience in order to further the narrative. One of the emotionally most impactful scenes is

---

3.   https://www.merriam-webster.com/dictionary/kinesthesia (last accessed 29-06-2021).

when I see Okita's family cross the border and then I am prompted to repeat that same step across a virtual line. It is a simple task that doesn't even need a hardware translation of the gesture. To grab something in VR, one needs to push a button on an input device. To take a step is to take a step with your own body. This gesture, combined with the symbolic meaning of the line, activates specific knowledge and/or memories stored in the body.

VR's special embodiment status lies in the hybrid and layered multisensory environment to which the viewer-explorer is exposed. The final experience is a result of the interplay of these stimuli coming from various physically existing and computational sources, creating a perceptually hybrid and multisensory environment. More than any other visual media, VR addresses the body in a particular way in which the body-self is experienced simultaneously as real and virtual. Kinesthesia and proprioception are the senses that situate the hybrid body-self in the enclosing environment and create the basis of the relationship between the viewer-explorer and the VR experience.

4   VR as Dream

Torben Grodal describes VR as a feeling of "strangeness and dreamlike un-realness" compared to film, where the viewer has a tacit knowledge that the film is a simulation. "This is because the closeness is 'our closeness,' and therefore it is more troublesome to be disregarded by the people close by in VR. [...] dream states are characterized by a hyper-activation of visual and acoustic perception whereas the muscles are totally immobilized, and thus active agency is absent in dreams" (Grodal 2018: 2). Although the body is not completely immobilized in VR and there is a certain amount of active agency, the interaction with the environment, characters and objects is not fully realistic. Objects do not have haptic resistance and characters cannot be touched or interacted with in the complex manner we do in our everyday life. VR has been compared not only to dreams but more generally to transcendental, shamanistic states (Jones 2000: 27), as a virtual space where dreams and desires can be projected. "Aesthetic illusion," a term used by Werner Wolf (2013), approaches the special state of mind in VR from the perspective of its aesthetic elements. He defines aesthetic illusion as "a specific imaginative, emotional and psychic response elicited by the reception of artefacts of various kinds, regardless of their aesthetic merits" (2).

Heterotopias offers an embodied experience that can be quite open-ended and abstract, compared to traditional action and character-based storytelling. Visually, the experience is highly stylized and rather dreamlike. The stereoscopic 360-degree recordings are color corrected and combined with computer-generated elements, so it is hard to separate the two sources. Besides the direct, bodily interaction, or rather immersion, through the hanging chair, the VR experience's main mode of interaction is utilized through blinking. Blinking, an involuntary action usually considered "noise" in the system, becomes the primary mode of interaction. The experience employs eye-tracking using the Fove headset to transform users' blinks into cuts. In *Heterotopias*, each blink triggers a new arrangement of visual and sonic elements. Consequently, space is continually redefined. The amplitude of the changes caused by blinking gradually increases during the experience, so that at some point (individually different) viewers become aware of the changes their blinking causes. By creating a feedback loop, the otherwise involuntary and unperceivable blink becomes apparent, conscious, and controllable. The user unconsciously-consciously performs the cognitive work of assembling the audiovisual experience, turning the concept of montage upside down (Kaplan and Ruszev 2018). The idea of blinking connects to the idea of the dreamlike state of VR. Although, physically, there is no blinking when we dream, the abrupt and unconsciously triggered shifts in the environment defy the continuity of perception familiar from our waking state. In this sense, the narrative continuity of the experience arises from a combination of an affective, bodily reflection and an abstract, symbolic reflection of the dreamlike experience.

5   Narrative in VR

Storytelling is a fundamental human activity that articulates, reflects, and shares our knowledge about the world. As the word itself contains an initial verbal character, narratology has long time focused on language-based storytelling. Modern narrative theories such as Russian formalism (Tynianov 2019; Šklovskij 2016) or

French structuralism (Genette 2002; Metz 1974) have been rooted in literature and therefore apply a linguistic approach. The rise of cinema as a new medium shifted storytelling from the domain of the written language into the visual realm. Nevertheless, cinema, as much as literature, has an author-driven narrative and can be connected to the two main Platonian categories of "diegesis" (the author directly addresses the audience) and "mimesis" (the author addresses the audience using characters) (Plato and Leroux 2013). VR and cinema are related as both operate in the visual realm. Yet, narrative in VR as compared to cinema is much less author-driven and prescribed. The viewer-explorer has the freedom to look around, move and, in some cases, interact within the VR experience.

Another aspect that falls under special consideration is the spatial and embodied character of narrative in VR. The viewer-explorer inhabits the virtual space with their body, even if their physical body does not exist within the virtual world. The senses of kinesthesia and proprioception are always activated in VR. The time and space of the virtual world become the real-time space for the viewer-explorer.

So what are some helpful ideas regarding narrative in other media that can be considered in VR? In his book *Poetics of Cinema* David Bordwell regards narrative as a "transmedium phenomenon" (Bordwell 2013: 3) and presents the elements that he considers to be essential to narrative. A narrative consists of events arranged in time including a continuity of agent and causal connections (4). Furthermore, a narrative involves some change that structures the narrative within an opening and ending point (6). Bordwell applies these ideas in the realm of the cinema and goes further in specifying other aspects of a narrative, driven by aesthetic elements of cinema.

This quite minimalistic and open definition of narrative can be useful in VR, as it points to the most fundamental structural elements of a narrative but leaves enough space for specific interpretations arising from VR as a medium. The questions are then how a narrative emerges from a participatory process and how it is contingent on VR's aesthetic specificities, such as an immersive, omnidirectional and spatial audiovisual environment and the embodied agency of the viewer-explorer? In this sense, we can talk about a story world, not as the totality of the "agents, circumstances and surroundings" (Bordwell 2013: 6), but as "the mentally constructed model of a 'universe'" (Hatavara, Hyvärinen et al. 2018: 1). Every VR experience constitutes its own universe that comes into being by the active participation of the viewer-explorer.

6   Cognitive approach: attentional theory and event segmentation

What are some ideas from the field of cognitive sciences studying perception in general and regarding film in particular that could be helpful in describing narrative in VR?

The "attentional theory of cinematic continuity" developed by Tim J. Smith studies the cognitive purpose of the rules of continuity editing. He looks at key elements of the continuity editing style, including match-action, matched exit/entrances, shot/reverse-shot, the 180-degree rule, and point-of-view editing. Smith explains the importance of the active role of the viewer in the "perceptual construction" of the film (Smith 2012: 2). He argues that the continuity "flow" of a film is dynamically constructed by the viewer, shaped by what the viewer is "attending to, what they are perceiving, and what they are expecting. [...] The continuity editing rules use natural attentional cues such as off-screen sounds, conversational turns, motion, gaze cues, and pointing gestures to trigger attentional shifts across cuts" (2). Furthermore, Smith describes the role of the saccadic movements and fixations of the eyes in perceptually stitching together the separate visual information and adding it to a working memory representation of the scene (6). In other words, the continuity of the visual environment is perceived in very short fragments during the fixations of the eyes and assembled into a continuous experience.

Another theory from the cognitive sciences regarding the perception of ongoing activity is the so-called "event segmentation theory" (Zacks, Speer et al. 2007), which assumes that the viewer constructs and maintains a mental representation of the current unfolding event and, based on this, they make predictions about possible continuations of the event.[4] When these possible predictions are violated, the viewer registers the beginning

---

4.   Here we have to note that the idea of mental representation and other embodied and enactive cognitive theories seem to be incompatible. The debate does not revolve around the question of whether or not we can segment our experience into separate events or

of a new event. On the side of film theorists, Bordwell and Thompson (2006) argued similarly in stating that continuity editing works by validating or refuting expectations. The most obvious violation of the expectations is the change of time and space in the next shot, which indicates the onset of a new scene. New events can also be perceived at a lower or higher level in film, perceiving a natural segmentation that constitutes the narrative flow. There can be a change within a scene, a turning point in the ongoing action, or a change in the emotionality or tonality of the scene. These changes can be punctuated by cuts as much as by other aural or visual cues – change of extradiegetic music, dialogues, acting, camera movement, change of lights and/or color. Event segmentation and the interplay between expectations and their violations play a role on a larger scale when making sense of the overall narrative structure of a film.

I argue that these two theses can be applied in the VR experience, in which the continuous environment is first fragmented through saccadic movements and segmented through expectation and prediction of future events. After this act of disintegration, the experience is reintegrated, based on the viewer's attention. Both these processes that are perpetually alternating and building on each other are key in "storifying" an experience.

7   Intersubjectivity, kinesthetic empathy, embodied simulation and the aesthetic elements of VR

Color, movement, camera angle, mise-en-scène, sound, and editing all impact the viewer on an affective, bodily level. Vivian Sobchack argues that the self-awareness of one's own embodiment is a radically irreducible condition of a reversible structure of empathy and sympathy for both the other body-subject and body-object, in what she calls intersubjectivity and interobjectivity, respectively. Her position on intersubjectivity allows for the "sanguine sense of not merely being-in-the-world but of also belonging to it" (Sobchack 2010: 246). This understanding of intersubjectivity gives us an interesting ground to grasp how the viewer-explorer might relate to their own virtual body on the one hand and how the feeling of belonging to a virtual world might result in the feeling of presence and empathy for other characters.

Tarja Laine draws a reciprocal relationship between film and spectator "among their bodily energy, affect, rhythm, valence and the very same attributes of the film's aesthetic system" (Laine 2018: 3). Moreover, the argumentative emphasis in her book *Feeling Cinema, Emotional Dynamics in Film Studies* is on the direct emotional engagement of the spectator with film aesthetics – albeit in a historically and culturally habituated fashion. Laine shifts the focus from a narrative character-driven analysis of a film toward the agential character of aesthetic elements that activates a perceptual, sensuous mental engagement by the viewer. Identification, empathy, or other reactions may arise but are not prescribed, as Laine argues (6).

I argue that this reciprocal relationship can be equally established between the aesthetic system of VR experience and the viewer-explorer. Instead of a linear succession of separate shots, a VR experience reveals itself as continuous space (or a succession of spaces). It therefore has a lower level of control over the attention of the viewer-explorer. Instead of cuts, a VR experience utilizes audiovisual clues that can guide the viewer-explorer's attention and shape their emotional and cognitive involvement and sense of narrative.

Change and repetition are basic actions that sustain attention cues. Color shift and movement, for example, affect lower-level feature perception. Change, affecting characters and objects, works on higher-level object perception. Changes that concern complex sequences are perceived as events. Cues can be aural such as a sound effect, a dialogue, or a piece of music. Cues can also have an interactive nature: objects that can be picked or interacted with, hotspots that can be looked at, movement within the environment itself that can be either a passive experience through the shifting space or active by moving virtually through space, or, if hardware allows, moving the viewer's own body in parallel to the movement in the virtual space. What is important to stress is that most of these clues/changes have an implicit embodied effect on the viewer that has to be calculated when designing an experience. The coordinate system of a VR experience is centered according to the body of the viewer. Things happen behind the viewer and not outside the film frame. Even cuts, dissolves, and wipes, when used, are experienced in an embodied way. All these elements constitute

---

if we can predict certain actions. In a simplified version, different positions within cognitive science debate whether this happens on a purely mental level or if and how the body is involved. See more in Menary (2010).

what Karen Pearlman calls "kinesthetic empathy" (Pearlman 2016: 7), which helps the viewer-explorer to align emotionally with the experience. In VR, other than in cinema this "kinesthetic empathy" can be achieved and enhanced through interaction by giving agency to the viewer-explorer within the virtual reality experience.

In a similar approach Vittorio Gallese and Michele Guerra investigate "embodied simulation" in the context of cinema with a short look at new media. They claim that part of the human neural resources that are involved in active interaction with the surrounding world can be reused for the mental process of perception and imagination (Gallese and Guerra 2020: 1). In other words, in order to understand the behavior of others, we rely on our own experiences and, based on those, we create an embodied mental simulation of what we encounter. This embodied simulation is the synthesis of what the viewer-explorer experiences in VR through their body. Gallese points out that neurons in the human brain show multimodal properties (182). This means that various senses are converging in our brain, that a tactile sense can be triggered by the image of a caressing hand or the description of the same image. The authors argue that this multimodality is primarily guided by our bodily experience and not by abstract mental construct (182). Although the book does not specifically apply embodied simulation to VR, it states that this direct relational modality can be utilized in any visual media to mentalize the behavior of other beings (200).

The VR experience *Book of Distance* is a good case study for the above-mentioned processes. The narrative space of the experience is built up, resembling the stages of a theater, where the viewer-explorer is guided through the episodes of the life of Okita's family. Stylized theater-like backgrounds and elegant use of light shape the progression of the narrative events. Combined with the presence of the author-narrator, this VR experience could have resembled the structure of a traditional documentary film. The added benefit of VR is the possibility of interaction. The viewer-explorer is invited to take in their hand the digital facsimile of family photos and newspaper excerpts. The fact that they can be moved close to the viewer-explorer's eyes and their indexicality enhances the emotional depth of the gesture. Repetition and mirroring are also gestures that are part of the narrative design and invite the viewer-explorer to actively involve their own body by using some objects bearing importance in the life of an immigrant. We can take a photo of what we see before the narrative progresses and in doing so mirror the act of retaining memories. These profound gestures add an emotional depth to a non-fictional narrative created in VR. They activate "kinesthetic empathy" on a deep bodily level and account for an engaging and stark VR experience.

8   Notes on Terminology

Finally, we can ask whether the term narrative can be applied to the medium of VR without modifications. Amendments regarding storytelling can be added either from the vantage point of VR being a spatial medium or the specificity of the embodied experience it creates. The term "ambient storytelling" has been coined by Jennifer Stein and Scott Fisher (2013) at the Mobile and Environmental Media Lab, exploring location-specific mobile storytelling as part of the Million Story Building Project. The project created an interface with the George Lucas Building at the University of Southern California using mobile phones, sensor networks, and software applications to create a responsive environment of collaborative storytelling. Although not used for VR, the term ambient storytelling covers well the spatial aspect of VR and could be well suited to cover new aspects of the narrative.

From the vantage point of embodiment, the term "somatic montage" (Waite 2016) introduces a new notion of cinematic montage for immersive cinema. According to Waite, "the concept of somatic montage addresses new forms of montage techniques that marry chronological with spatial sequencing into an embodied, participatory creation of narrative" (2). The term somatic montage addresses both the "distribution" of images in the (virtual) space and the active embodied relationship between the viewer and the immersive experience (3). Both terms hold some truth in addressing the relationship between the viewer, the space and the narrative. While "ambient storytelling" highlights the spatiality of VR, "somatic montage" pertains to the viewer's active and embodied participation in the construction of the experience.

"Embodied narrative" could be a third possible term (Menary 2008). It has been used in cognitive sciences to point at simulative thinking processes seen in the context of the brain-body-world interaction. These ideas go back to Lakoff and Johnson's (1999) claims that the cognitive structures assisting thinking and language

making arise from the organism-environment interaction and go as far as to claim that the self is narratively constructed (Dennett 1991; Velleman 2006). Without the claim that any of these terms have to be canonized, I call for a further investigation of the terminology of the narrative in VR in order to be able to capture all its specificity and particularity.

9   Conclusion

Based on the examples and theoretical implications, we can generally say that the narrative experience of VR is inherently embodied and spatial. Narrative in VR is a process, emerging from a story world and shaped by the participatory and embodied attention of the viewer-explorer. As the viewer-explorer is positioned at the center of the VR world, we have to investigate their involvement utilizing ideas from the fields of cognitive sciences and cinema and media studies in order to account for the embodied nature of this relationship. The viewer-explorer enters a story world and creates a unique narrative based lived through their body. There is still a great potential in VR as an emerging medium to develop a new storytelling language that needs a novel approach in theorizing it. This article is an attempt to accumulate interdisciplinary theories and ideas that help reflect narrative in VR.

References

Aylett, Ruth and Sandy Louchart (2003). "Towards a Narrative Theory of Virtual Reality." *Virtual Reality* 7(1): 2-9. https://doi.org/10.1007/s10055-003-0114-9.

Bastian, Henry (1880). *The Brain as an Organ of Mind.* London: Keagan Paul.

Blach, Roland (2008). "Virtual Reality Technology - An Overview." In *Product Engineering: Tools and Methods Based on Virtual Reality*, edited by Doru Talaba and Angelos Amditis, 21-64. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-8200-9_2.

Blanke, Olaf and Thomas Metzinger (2003). "Full-body Illusions and Minimal Phenomenal Selfhood." *Trends in Cognitive Sciences* 13(1): 7-13. https://doi.org/10.1016/j.tics.2008.10.003.

Boal, Augusto (1999). *Legislative Theatre: Using Performance to Make Politics.* London: Routledge.

Bordwell, David (2013). *Poetics of Cinema*. New York: Routledge.

Bordwell, David and Kristin Thompson (2006). *Film Art: An Introduction*. New York: McGraw-Hill.

Boring, Edwin (2019). *Sensation and Perception in the History of Experimental Psychology*. Delhi: Fascimile Publisher.

Chemero, Anthony (2009). *Radical Embodied Cognitive Science.* Cambridge, MA - London: The MIT Press.

D'Aloia, Adriano and Ruggero Eugeni, eds. (2014). *Neurofilmology: Audiovisual Studies and the Challenge of Neuroscience*. *Cinéma&Cie. International Film Studies Journal* XIV(22-23).

Foucault, Michael (1986). "Of Other Spaces: Utopias and Heterotopias". *Diacritics* 16(1): 22-27.

Gallese, Vittorio and Michele Guerra (2020). *The Empathic Screen: Cinema and Neuroscience*. Oxford: Oxford University Press. https://dx.doi.org/10.1093/oso/9780198793533.003.0004.

Genette, Gérard (2002). *Figures*. Paris: Seuil.

Grau, Oliver (1999). "Into the Belly of the Image: Historical Aspects of Virtual Reality." *Leonardo* 32(5): 365-371.

Graziano, Michael and Matthew Botvinick (2002). "How the Brain Represents the Body: Insights from Neurophysiology and Psychology." In *Mechanisms in Perception and Action: Attention and Performance*, edited by Wolfgang Prinz and Bernhard Hommel, 136-157. Oxford: Oxford University Press.

Grodal, Torben (2018). "Virtual Reality Experiences, Brain, Body, and Muscular Agency - An Embodied Approach Informed by Neuroscience." Preprint.

Hatavara, Mari, Matti Hyvärinen, Maria Mäkelä and Frans Mäyrä (2018). *Narrative Theory, Literature, and New Media: Narrative Minds and Virtual Worlds*. New York - London: Routledge.

Jones, Stephen (2000). "Towards a Philosophy of Virtual Reality: Issues Implicit in Consciousness Reframed." *Leonardo* 33(2): 125-132.

Kaplan, Noa and Szilvia Ruszev (2018). "Heterotopias—Optical Mastication and Spatial Reconfiguration." In *Artistic Research Will Eat Itself*, The 9th SAR International Conference on Artistic Research University of Plymouth, April 11th-13th, edited by Geoff Cox, Hannah Dreyson et al, 197-203.

Kilteni, Konstantina, Raphaela Groten and Mel Slater (2018). "The Sense of Embodiment in Virtual Reality." *Presence: Teleoperators and Virtual Environments* 21(4): 373-387.

Laine, Tarja (2018). *Feeling Cinema: Emotional Dynamics in Film Studies*. New York: Bloomsbury.

Lanier, Jaron (2017). *Dawn of the New Everything. Encounters with Reality and Virtual Reality*. London: Vintage.

Mazuryk, Tomasz and Micahel Gervautz (1996). "Virtual Reality History, Applications, Technology and Future." Institute of Computer Graphics, Vienna University of Technology.

Menary, Richard (2008). "Embodied Narratives." *Journal of Consciousness Studies* 15(6): 63-84.

Menary, Richard (2010). "Introduction to the special issue on 4E cognition." In *Phenom Cogn Sci* 9: 459-463. https://doi.org/10.1007/s11097-010-9187-6.

Metz, Christian (1974). *Film Language. A Semiotics of the Cinema*. New York: Oxford University Press.

Pearlman, Karen (2009). *Cutting Rhythms. Shaping the Film Edit*. New York: Focal Press.

Plato and Georges Leroux (2013). *La République*. GF Texte integral 653. Paris: Flammarion.

Popat Sita (2016). "Missing in Action: Embodied Experience and Virtual Reality." *Theatre Journal* 68(3): 357-378.

Ryan, Mary-Laure (2001). *Narrative as Virtual Reality*. Baltimore: Johns Hopkins University Press.

Shklovsky, Viktor and Aleksandra Berlina (2016). *Viktor Shklovsky: A Reader*. New York: Bloomsbury.

Smith, Tim (2012). "The Attentional Theory of Cinematic Continuity." *Projections* 6(1): 1-27. https://doi.org/10.3167/proj.2012.060102.

Sobchack, Vivian (2010). *Carnal Thoughts. Embodiment and Moving Image Culture.* Berkeley: University of California Press.

Stein, Jane and Scott Fisher (2013). "Ambient Storytelling Experiences and Applications for Interactive Architecture." AMBIENT 2013: The Third International Conference on Ambient Computing, Applications, Services and Technologies.

Tynianov, Yuri (2019). *Permanent Evolution: Selected Essays on Literature, Theory and Film*. Boston: Academic Studies Press.

Waite, Clea (2016). "Somatic Montage for Immersive Cinema." In *Organization in Early Soviet Thought: Bogdanov, Eisenstein, and the Proletkult*, edited by Pia Tikka. Helsinki: Aalto University.

Wolf, Werner, and Walter Bernhart, eds. (2013). *Immersion and Distance: Aesthetic Illusion in Literature and Other Media*. Amsterdam: Rodopi.

Zacks, Jeffrey and Khena Swallow (2007). "Event Segmentation." *Current Directions in Psychological Science* 16(2): 80-84. https://doi.org/10.1111/j.1467-8721.2007.00480.x.

Szilvia Ruszev – University of Southern California (US)

✉ ruszev@usc.edu

Szilvia Ruszev is a film editor, media artist and scholar working across different media formats. Her broader research interest focuses on sensuous knowledge, montage theories and politics of post cinema. Her own artistic work relates to very personal moments, certain states of emotional solitude in relation to the Other, both in its particular and abstract notion. As editor, she collaborated with internationally acclaimed directors such as Peter Greenaway, Anders Østergaard, and János Szász. Her award-winning work has been part of numerous international film festivals and exhibitions such as Karlovy Vary IFF, TIFF Toronto, Berlin IFF, Siggraph and Codame. Between 2010-2016, she was a faculty member of the Editing Department at the Film University Babelsberg Konrad Wolf. Currently, she is an Annenberg Fellow, pursuing a Ph. D. degree in Media Arts + Practice in the School of Cinematic Arts at the University of Southern California.