Jeff Rush
Department of Film and Media Arts
Temple University
215-204-4372
jrush@temple.edu

MiT4: Media in Transition
May, 2005

The Ergodic Bridge

In his oft-cited book, *Cybertext*, Espen Aarseth introduces the term "ergodic" to describe the reception of interactive discourse. Ergodic, a term appropriated from physics, "derives from the Greek words *ergon* and *hodos*, meaning 'work' and 'path.'"(Aarseth *Cybertext* 1). It describes a "type of discourse whose signs emerge as a path produced by a non-trivial element of work" (Aarseth "Aporia and Epiphany" 32). Ergodic discourse includes the *I Ching*, hypertext, interactive media, computer games, automated poetry generators, and Multiuser Discourse (MUD's) among others. Aarseth's work is sometimes described as moving the study of interactive texts beyond literary, dramatic or film theory to that of cybernetic systems (Frasca 23).

Aarseth defines cybertext as text that has an information feedback loop which "centers attention on the consumer, or user, of the text, as a more integrated figure than even reader-response theorists would claim."(*Cybertext* 1) The ergodic is the process of engaging with cybertext, an extranoematic performance that goes beyond what takes place in the reader's head. Cybertext is never complete. "You are constantly reminded of inaccessible strategies and paths

not taken, voices not heard." (*Cybertext* 3) While traditional critical methodologies have some role to play in considering cybertext, "none of these have expressed the perspective of the text as a material machine, a device capable of manipulating itself as well as the reader."(*Cybertext* 24)

In building a typology of cybertextual strategies, Aarseth first defines "textons" the strings of signs that appear in the text and "scriptons" the strings of signs extracted by the reader, along with a traversal function – "the mechanism by which scriptons are revealed or generated from textons and presented to the user of the text." (*Cybertext* 62) Then he categorizes cybertexts by evaluating them against seven variables: dynamics—the permanence of scripton contents; determinability—the stability of adjacent scriptons; transiency—the stability of scriptons over time; perspective—the user's potential strategic role as a character; access—the user's degree of random access to the text; linking—the traversal choices presented to the user; and user functions—the user's opportunity to add textons and traversal functions to the text.

In abstracting these variables, Aarseth speaks to the manner in which textons are organized and presented by the interactive program, and how the user may select textons to become scriptons. This positions interactive media outside the traditional literary categories and directs our attention in an important way to the tension between user choice and machine design. Since we would expect that much of the expressive power of interactive media lies in the formal structure of its transversal interactivity, this analysis proves to be a major step forward.

But the user's extracting of scriptons from textons does not take place in a contextual or expressive vacuum. The transversal function, so critical to defining the texture of interactivity, must also express a broader meaning than merely its style of linking if interactivity is to be more than a modeling of physical process. What Aarseth's model does not address is the pleasure we may find in the content of interactivity, or the world it may uniquely open for us. For instance,

he rules out narrative. Aarseth concludes *Cybertext* with a plea to get out from beneath a dystopia, “the omnipresent influence of narrative, both as hegemonic theories of discourse and as a socially dominating aesthetic mode.”(*Cybertext* 182) Later he says approvingly referring to computer games, “there is no such thing as the unfolding of a predetermined story”("Aporia and Epiphany" 35). Aarseth continues this attack in his response to Janet Murray’s argument that stories and games be recombined (Murray). “Games will be games and gamers will be gamers. Storytelling, on the other hand, still seems eminently suited to a sequential formats such as books, films, and e-mails . . . ” (Aarseth "On Line Response" 10)

Aarseth is hardly the only scholar to raise questions about the role of narrative in interactive media. Although not as dismissive, Lev Manovich sees narrative as subordinate to the database as the organizing principle of New Media. “In the database/narrative pair, database is the unmarked term,” (228) with “unmarked” referring to the more general, the more likely to be given. Narrative is merely one of a number of ways to interact with the database. Marie-Laure Ryan identifies two myths tied to the computer and narrative – the myth of Aleph in which the computer can generate an infinite number of stories out of a finite text, and the myth of the Holodeck in which the user becomes a character in a Shakespearean-like drama – both of which she calls into question. (Ryan) Yet Aarseth’s rejection of narration along with his claim that ergodic art has “no such thing as the unfolding of a predetermined story” are more pointed than arguments made by either Manovich or Ryan in bracketing interactive media not only from story, but from a narrative frame.

I support Aarseth’s position that there must be a cybernetic dimension to all analysis of interactive media. Interactivity is about machine logic, process and control, the interplay between noise and information. However, as Mark Hansen explains when discussing Donald

MacKay's critique of Claude Shannon's context-less information model, MacKay is "concerned with reconciling *two* processes [italics mine], or two sides of the process of communication: on the one hand, the production of representations, and on the other, the effect or function of representations, which is equivalent, as we shall see, to their reception. . . . " (77) The two sides of the representations I will explore in this study belong to that particular subset of digital media that consists of interactive narrative, more specifically mixed-media interactive narrative. My approach will combine the ergodic along with a number of angles on narrative pragmatics including traditional narrative theory, rhetorical narrative theory, Bergsonian takes on intensive/extensive time, and cognitive/systems theory.

First, I will need to rethink the ergodic. While the model that Aarseth proposes is a major advance toward understanding interactive narrative, it only accounts for one half of the ergodic experience – either the users' interaction with computer or the representation of this on the screen, depending on which section of Aarseth we read. I will argue for a more robust model, an ergodic bridge, that ties user action to representation, interactivity to narration.

* * *

In *Cybertext*, Aarseth focuses on ergodic literature and how its reception differs from that of other forms of reading. Since my concentration is on mixed-media interactivity, I will initially approach the ergodic through a subsequent Aarseth article, "Aporia and Epiphany in *Doom* and *The Speaking Clock*: The Temporality of Ergodic Art", where he broadens his model to include computer games. Although my focus is not on computer games per se, Aarseth's treatment of *Doom* provides an excellent statement of his argument as it applies to audio-visual media.

Aarseth draws on the distinction between description and narration made by narrative theorist Gerard Genette in a 1966 paper entitled “Frontiers of Narrative”. Genette defines description as the representation of objects or characters in a timeless state (“the house was white, with a red door”), and narration as the representation of actions and events as they unfold in time (“the student pulled a knife and stabbed the professor”). Aarseth builds on this by claiming first that traditional narrative contains both description and narration, but no user action or “ergodic elements”. Next, he notes that non-textual games, he cites soccer[1], contain only actions which he calls “ergodic elements”, but neither description nor narration. Finally, he argues that interactive art and computer games contain both ergodic elements and description, but no narration. Thus, no narrator. In having no narrator, or more broadly, no narrative agency of any sort, ergodic art does not contain any form of narrative. According to Aarseth, the reason that computer games have no narration is that they have “no such thing as the unfolding of a predetermined story” or put another way “the event space is not fixed before the time of play” ("Aporia and Epiphany" 35).

---

[1] Actually he cites football and we need to distinguish the narrative elements in football as it is known in the United States as opposed to football in the rest of the world. In international football, soccer, action is indeed continuous and arguably not narrated by the player on the field. American football, however, is quite different. What takes place in the football huddle is discursive, being a coded message which the other players are equipped to decode, and narration, a representation of action, even if it is yet to come. From this narration, we can construct the action (at least if all goes as planned).

One more digression. Aarseth’s argument that there is no narration in a sport like football is restricting the definition of narration to physical action alone. But all sports have rules that define what constitutes a score, and what can or not be done between the players, along with fields or courts that define the boundaries of play. While not strictly narration, this combination of rules and field definition does serve to narrate, or shape, the actions that take place. It is because the player has internalized these narrative constraints that she is able to play.

Further, our culture narrates games through the media so that if we follow them we create certain models, or maybe fantasies, for ourselves, of great plays that we’d like to make. Not only do many of us who grow up in the United States for instance internalize the fantasy of the bases-loaded home run or the winning three-point play, but the characteristic shape of sporting events. This is why a truly outstanding moment is not only a matter of athletic performance that can be appreciated by everyone, but an historical aberration that can only be understand by someone who has immersed herself in the culture of the sport. The Red Sox beat the Yankees.

The claim that the ergodic does not contain any form of narration raises an important question about just what Aarseth means by the term. He argues that the computer game *Doom* creates its world through both ergodic action and description. Yet ergodic action, which he compares to playing an athletic event, would seem to reside outside the text, leaving description as the only textual element contained within *Doom*. Alone, description can not generate the story movement necessary to create narrative. As Genette puts it, "description might be conceived independently of narration, but in fact it is never found in a so to speak free state. . . " (134). So, if *Doom* were solely made up of extra-textual ergodic action and textual description, it would remain textually static, inert. Aarseth would be right – *Doom* would lack narrative. But it would also not play. There would be nothing on the screen except background.

But does ergodic action reside outside the text as Aarseth's athletic metaphor would seem to suggest? Or is ergodic action really the textual representation of that athletic metaphor, the image on the screen that translates the mouse move or keyboard click? Or is it both? Aarseth is unclear. Sometimes he defines ergodic action as analogous to the non-textual playing of a sporting event, direct, non-mediated and outside of a text. Sometimes he states that ergodic action refers to the "path of signs" produced by work which would seem to suggest that the signs represent action in the fictional world, presumably triggered by the user.

My reading is that Aarseth approaches these two elements – the user's action and its translation on the screen – as part of one larger concept and that this unity constitutes the cybernetic thrust of his argument, its focus on control. Extra-textual work leads to a "path of signs", ergodic action accounts for both the control and the screen-based action of *Doom*. *Doom* has "no such thing as the unfolding of a predetermined story" because it is created anew each time by the user. As Aarseth notes, "When a system is sufficiently complex, it will by intention,

fault, or coincidence, inevitably produce results that could not be predicted even by the system designer" (*Cybertext* 27).

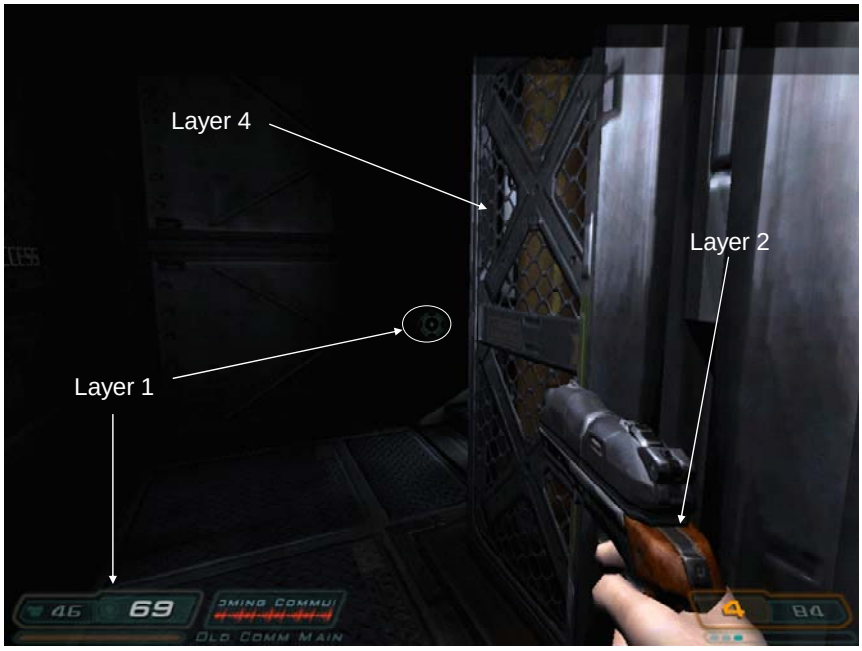
Aarseth's fusing of user action and its representation, however, glosses over two related, but distinctly different functions. Implicit in his model is that there is only one to represent the user's action on the screen. Or that there are several ways, but there is little distinction between them. However, this would only hold true if the image on the screen were an exact replication of the user's hand as it moved the mouse in real time, and, of course, this too would be a representation, a textually remaking that would have to account for the translation from three to two dimensions, different light balances, a static camera position and a lack of user sensation among other issues.

The distinction between control and representation that Aarseth seems to elide will be central to this study. To keep it untangled, let's call the extra-textual action taken by the user "ergodic[direct]" and the textual action that occurs as a result of user action "ergodic[represented]". The translation between these two perspectives, the phenomenon that makes interactive media unique by challenging the boundary of the text, we will call the "ergodic bridge".

* * *

Although not the first, *Doom*, which was released in late 1993, was the most popular of the original generation of the "first-person shooters". This genre projects the literal point of view of a character, generally signified by a foreshorten hand waving a gun, into the screen's depth so that we move with the character through a world where we must shoot or otherwise eliminate monsters out to destroy us. It can be so visceral and haptic that we may flinch as monsters or

missiles appear to fly out of the screen at us. Below is a scene from *Doom3*[2] which we will use to illustrate the ergodic bridge.



The foreground-most layer, I'll call it layer one, shows blocks of text listing the score, the numbers of bullets in the gun, and the location. It also contains a circle in the exact center of the screen which represents the gun's aiming point. This layer does not shift relative to the edge of the screen when I pan or move the character. Layer two contains the weapon itself, in this case a hand gun, along with three fingers and the top part of a hand. When the point-of-view character remains still, the gun seems to float up and down, but the circle which represents its aim does not move. The gun is markedly foreshortened so that the circumference of the barrel further from us is half that of the barrel nearer. The hand too is foreshortened with the tip of the forefinger markedly narrowed relative to the wrist in the foreground. For the moment, I see no third layer.

[2] Here I'm going to cheat. Aarseth used the original *Doom*. In this paper, I'm going to use the newer *Doom 3* which functions better on a contemporary computer. I acknowledge that more contemporary computer games tend to increasingly feature narrative elements. However, I take Aarseth's argument to be broader than turning on the simple issue of technical refinement so I do not feel my choice of a later version (which incidentally is a recasting of the same "story" as the original *Doom*) undercuts my argument.

The final and fourth layer contains the environment itself, in this case lacking any motion. When motion is displayed here, it is iterative, such as dials moving back and forth, or lights sweeping the scene.

Before I apply any ergodic action, then, I am aware of a still background and a foreground with only the iterative motion of the gun floating up and down. This clearly meets Aarseth and Genette's definition of description. In fact, going further into Genette's article, it meets not only his criteria for the first order of description as ornament, but the second, deeper function which defines description as exposition because it suggest directions in which the story might advance. Beyond this, nothing happens unless I take the first ergodic[direct] step, manipulating the mouse or clicking the keyboard. In this case, I click the keyboard command 'w' which moves me forward. Layers one and two do not move. Rather the descriptive layer four appears to advance toward me and I hear "my" footsteps creating the illusion of walking. The third layer which I had not previously been able to identify suddenly springs into view as a character (actually a Zombie) attacking me.

Seeing the Zombie, I push the mouse (ergodic[direct]). The background and the Zombie shift, creating the illusion that I have moved relative to them (ergodic[represented]). Creating the illusion of movement by adjusting the background and the Zombie as opposed to moving the point-of-view character is common to most cell animation. However, although it is common, it has the effect of creating the illusion that the character I control is the center of the fictive world which adjusts itself accordingly to meet my needs. This perspective is further heightened because the world is seen entirely through my character's point of view. The combination of how the movement is created and how it is shown, both expressions or properties of ergodic[represented], reinforces my feeling of power. This marks a perspective, a construction of a view on this world, a narrative act.

This narrative act is what Aarseth rejects. For instance, in demonstrating the lack of narrative act in *Doom*, he identifies as narration his *recounting of the playing* of the game in contradistinction to the game itself. He says, "I base my narrative [of game playing] on a more

fundamental structure, the event space of aporias and epiphanies, which are the prenarrative master-figures of experience, from which narratives are spun." (Aarseth "Aporia and Epiphany" 39) The aporias he mentions are those he discovers in *Doom*, including one where confronted with too many monsters to kill individually, he realizes it is necessary to shoot barrels of toxic waste whose explosions will eliminate the monsters collectively. The epiphanies come as a result of successes in completing such puzzles. By saying that this pattern of aporia and epiphany in *Doom* represents "prenarrative master-figures of experience", as though they either happened by accident or represented some culturally and cognitively-autonomous, natural shaping of experience, Aarseth argues that the game and the narrative act which organizes it are independent of mediation, meaning or narration.

But all ergodic[represented] acts have meaning; they all represent a shaping of the game world. For instance, when I roll the mouse, the screen rotates around the fixed axis of my position creating the illusion is that I am looking without moving. However, my look is represented as passive. My eyes are the receivers of stimulus rather than triggers for it. My look never challenges or provokes action. This suggests that I am projected as a disembodied presence into the story world, where my only agency is my gun. I can observe without influencing the system I am observing. The same disembodiment occurs when I walk. I can run into things and knock them over. But although my implied body has mass, I can only see the damage it causes, not the body part that is responsible. I am placed in the Z-axis so that all the game activity comes directly towards the user, implying a way of recounting the world that emphasizes danger, haptic involvement and immediacy. My inability to see without fists or guns in the foreground is a narrative act which refuses to represent whatever other qualities my character might have in favor of the brutal. The fixed position of the aiming circle in the center

of the screen functions in the same way. Whether I am shooting or not, the act of aiming becomes the central metaphor around which the world is constructed.

We can identify further narrative acts. As we know from art history, the representation of depth is a construction rather than purely an imitation of the physiology of the eye[3]. While not so extreme that I feel the abstraction, the use of the exaggerated depth draws me into the story world. The perspective is frequently two-point; one vanishing point being in the direction in which I am headed and the second being another path I might take, graphically illustrating the choice of direction I constantly face. Although there are commands that move me sideways, the exaggerated depth, as well as the design of the obstacles I face, makes such movements of limited attractiveness. The extreme lighting contrast exaggerates the unknown and makes me curious, or frightened, to enter it. The fact that the character has no effect on the story world by virtue of his looking or the noise he makes walking increases the narrative single mindedness, as does his lack of speech. The contrast between my lack of embodiment set against the stark concrete, presence of the gun furthers this focus.

The images and sounds that ergodic[represented] display are triggered by ergodic[direct] action, but their representation is beyond my control and governed by a set of rules that predetermine the events I will face. My ability to impose my will is quite limited. I can decide where to move and when to shoot, but, without reprogramming the game, I can not rearrange the way my actions are reflected back to me from the fictional world. I can not, for instance, make my viewpoint into the screen independent of where the gun points. I can not choose to look sideways as I move forward to watch my arm swinging at my side. And I can not avoid the

[3] Of course, we know now that the eye doesn't transmit "reality" either, but rather shapes its message to the needs of the organism.

puzzles that are pre-planed for me, without taking the shortcuts that I want to think I have discovered, but which are really designed into the fabric of the game.

Thus, while *Doom* provides many structured items that the viewer can control, the game is not free of narration or narrative agency as Aarseth suggests. Rather it is constructed around the tension between cybernetic control and narrative pragmatics with the ergodic bridge the function that connects them. I will go into greater detail about the ergodic bridge later in this study, but suffice it to say now that it is the prime site of expressivity in interactive narrative, the core of its narrative agency. As in literature and in film, a shift in the representation of narrative agency will change the feel of the game. We can see this more clearly if we contrast *Doom* to a later subgenre of shooter games, the tactical first person shooter, in this case a game called *Rainbow Six: Raven Shield*.

Raven Shield requires me to track down terrorists by "commanding" two squads of four fighters each. Unlike *Doom*, in *Raven Shield* I can move between and play each character, while giving simple orders through keyboard command to the others in my squadron. When playing one character of course I can see the other characters around me; should I die as the point of view character/narrator, the game does not stop, but rather my point of view goes to ground and I am given an option of jumping to a more fortunate character. Unlike *Doom* where there is no preplanning, *Raven Shield* requires a great deal of set up where I am briefed on my mission, have to choose a team and study a tactical plan. Team members are given a number of variables such as leadership, self-control and demolition.

Doom presents an almost perfectly disembodied, first-person narration. The only opportunity I have to see the character I play is during cut scenes, non-interactive, animated sequences where my character becomes one more actor on the screen. In addition to being third-

person, the cut scenes tend to be shot and edited with such omniscience that not only am I outside looking in, but I have to work to identify my character on the screen. *Raven Shield* adds another dimension. When I am playing one character, I still can not see any part of myself except for a section of forearm and my rifle. However by shifting to another character/narrator, I can see myself from outside and watch myself move or react. This does not work for very long. I quickly forget the position I occupied inside the point-of-view of the first character and take the point-of-view of the second, becoming disembodied once again. However, during that moment when I feel I am now watching the character whose eyes I was just viewing through, I find myself having an unexpected sense of identification. I experience the dual impression of seeing from the inside and the outside I associate with close-in, third-person fiction; although masked and armed, the character I once was seems frailer than I expected, vulnerable as he huddles with his fellow soldiers, because I project onto his image the susceptibility I felt when I was responsible for his point of view.

In the *Raven Shield* screen still below, the ergodic bridge communicates a very difference sense than it does in *Doom*, at least as I perceive it during that moment after I have changed character/narrators. Because I have to reorient myself spatially in the represented world, my confidence in the ergodic[direct] slips. I have to think consciously of the cognitive connection between the way I move the mouse and the way that movement is represented on the screen. This changes my perception of the ergodic[represented] so that instead of reading the screen as a realistic representation of pathways through which I might proceed, I find it initially fragmented into abstract shapes and then coherent, but disorienting. This alters my transactional relationship with the game. Before switching character/narrators, the deal I made with the narrative representation was that I would act upon it; my agency would drive it forward. Now in this

moment of reorientation, I instead query it, seeking clues as to which way I am looking. This confusion over direction has the effect of de-temporalizing space—that is, I not longer see space as representing the past where I have been and the future where I will go, but as signifying static timelessness and disorientation. Abstracted in this way, it becomes more figural or extensive than it normally would when my goal is simply to pass through it, providing a momentary metaphorical snapshot rather than the normal metonymic progression that drives narrative. This lack of confidence in its simple representation creates a clarity of vision that I think is characteristic of interactive narrative; I feel like I am on the edge of something that I see sharply, but that I do not quite understand, before it resolves into yet another navigational choice.

Looked at other way, the movement from character/narrator to character/narrator has at least the potential of changing the cut between observer and narrative system, although this does not actually happen in *Raven's Shield* since none of the character/narrators question the importance of the mission and hence can not jump out of the narrative system they occupy. In a more narratively complex piece, the jump between character/narrator might also be a jump out of an observer's position with the subsequent reinsertion into another character/narrative suggesting a remaking of the observer's position and a reconceptualizing of the narrative system. I can envision an interactive narrative that constructs such changes of observer's position in order to account for the blind spots inherent in that of the observers who went before.

Finally, there is the purely narrative change between *Doom* and *Raven Shield*. The fact that I can enter each of my teams' heads creates a sense of shared subjectivity, a deeper feel of community. My ability to see myself (although this keeps changing) both from the inside and the outside personalizes my relationship to the story world more so than I experienced with *Doom*. While the purpose of *Raven Shield* is still to kill bad guys, the ergodic bridge has shifted

its focus slightly, increasing my interest in my teammates and building a greater narrative into what is still functionally a shooter.

* * *

In this paper, I have proposed a revision of Aarseth's ergodic that splits his original model into two: the ergodic[direct] which identifies the user's manipulation of the computer's inputs and the ergodic[represented] which identifies the translation of that manipulation onto the screen. Neither perspective is a pure form in that they constantly feed back on one another. As contested as they are by the very feedback that ties them so fundamentally to interactivity, these two viewpoints provide an analytic model that will allow me to approach ever more complex, mix-media interactive narratives from both a perspective of control and data flow, and of narrative pragmatics. My goal for the rest of this study will be to explore how these two methodologies transform themselves as they feed back on one another.

Works Cited

Aarseth, Espen. "Aporia and Epiphany in *Doom* and *the Speaking Clock*: The Temporality of Ergotic Art." *Cyberspace Textuality: Computer Technology and Literary Theory*. Ed. Marie-Laure Ryan. Bloomington: Indiana University Press, 1999. 31-41.

---. *Cybertext: Perspectives on Ergodic Literature*. Baltimore, Md.: Johns Hopkins University Press, 1997.

---. "From Espen Aarseth's Online Response." *First Person: New Media as Story, Performance, and Game*. Eds. Noah Wardrip-Fruin and Pat Harrigan. Cambridge, Mass., London: MIT Press, 2004. xiii, 331 p.

Frasca, Gonzalo. "Simulation Versus Narrative: Introduction to Ludology." *The Video Game Theory Reader*. Eds. Mark J. P. Wolf and Bernard Perron. New York: Routledge, 2003. xxii, 343 p.

Genette, Gérard. *Figures of Literary Discourse*. European Perspectives. New York: Columbia University Press, 1982.

Hansen, Mark B. N. *New Philosophy for a New Media*. Cambridge, Mass.: MIT Press, 2004.

Manovich, Lev. *The Language of New Media*. Leonardo: Leonardo (Series) (Cambridge, Mass.). Cambridge, Mass., London: MIT Press, 2001.

Murray, Janet. "From Game-Story to Cyberdrama." *First Person: New Media as Story, Performance, and Game*. Eds. Noah Wardrip-Fruin and Pat Harrigan. Cambridge, Mass., London: MIT Press, 2004. xiii, 331 p.

Ryan, Marie-Laure. "Beyond Myth and Metaphor: Narrative in Digital Media." Poetics Today 23.4 (2002): 581-609.