

Dorit Naaman  
Queen's University

## Minding the Gap: Visual Perception and Cinematic Gap Filling

### Introduction

Imagine a screening of *Vertigo* stopped once Madeline is found dead by the tower, and the audience is asked "What happened? How did she die? Was she killed?" The audience may find it difficult to answer definitively; different perceivers would provide different scenarios, and they are unlikely to argue about which is true, but instead assume that the film will provide both explanation and closure in due time. Indeed, narratives operate in a curious way: while they tell us stories, we, the perceivers, are rarely willing to commit to plot lines or even to predictions about the progression and conclusion of the story before the text delivery is over. While narratives set-up expectations (which take the form of hypotheses), they often take new and surprising plot directions, ones that require perceivers to rearrange knowledge of plot in significant ways. Importantly, then, perceivers are ready to alter, cancel, or embrace new hypotheses as the text provides them with new information. In other words, the *fabula*, or the complete story is a product that a perceiver commits to only after the perception of the text is over.<sup>1</sup> And while the narrative as a product is being constructed during the perception, it is constantly in flux, or open to be in flux, until perception is over. Consequently, the conclusive narrative of a text is a post-perception product. Moreover, as a post-perception product, the narrative is constructed from memories reorganized in a causal order so as to yield the most coherent story possible.

Thus it is important to understand that story, or *fabula*, is a product of an array of high-order cognitive activities significantly different from low-order perceptual processes.<sup>2</sup> Empirical research on narrative suggests that narrative structures are a product of high-order mental operations. In a series of related experiments, Gee and Grosjean asked subjects to read and then recount a short narrative. They analyzed spontaneous pause duration between sentences and then matched them with Lehnert's complex analysis of narrative structure into simple plot units. What Gee and Grosjean found was that "as the narrative complexity of a break between two sentences increases, the pause produced by a speaker also increases—and in a very systematic way" (72). But while Gee and Grosjean were primarily interested in providing empirical evidence of narrative structure, their research reveals

another important phenomenon. They found a correlation between story parsing and pauses only when subjects retold a story after reading it to themselves. When they asked subjects to read the story aloud (even in a second reading), the pauses did not equally well match the narrative structure parsing (81). Gee and Grosjean do not explain why spontaneous retelling reflects so much more accurately a story's narrative structure, but in the context of our discussion it is clear; retelling takes into account that a narrative has been fully comprehended and interpreted before it is retold. Because it is a post-perception activity, the performer in the retelling (the former perceiver) has a full concept of the narrative. Reading out loud, by contrast, does not provide the correlation between pauses and narrative complexity, because it is hard for the perceiving agent to anticipate accurately where narrative units begin and end (or where pauses should be placed). According to this explanation, the correlation between retelling and narrative complexity is easily understood, as both require a post-perception moment in order to be fully actualized.

In this explanation of narrative, gaps configure in interesting ways. They provide a unique moment inviting perceivers to produce a hypothesis, but they also remind them that they may be wrong or that the text may be misleading them. The perceiver of a text, upon encountering a narratorial gap, is required to devise a strategy by which to fill the gap and make the text cohere. Such a strategy generally involves coming up with hypotheses as to what is likely to have happened, hypotheses the perceiver hopes are correct. In logical terms, the process of coming up with inferences to the best explanations is called abductive reasoning. Abductive reasoning not only deduces a set of possible worlds that could exist based on the facts we have (i.e., in our case different narratorial hypotheses) and on axioms, but also provides a way to codify the preference of one model over another. But while watching a film we are rarely aware of how or why we prefer a certain hypothesis over others. As I will show, the reason for this lack of awareness has to do with the differences between the processing of visual versus verbal information.

It is only after the transmission of the text is over that the perceiver can reorganize all the data, determinately fill in some gaps, and confidently claim that others shall remain open. But saying that a narrative (as a product) is a post-perception result does not mean that the construction of the narrative (as a process) happens entirely in post perception or in high-order cognitive arenas. As I will show, a narrative is constructed as perception occurs, and it is constructed based on bottom-up perceptual activities, as much as by top-down impositions of belief, ideology, and expectations.

### Theorizing Gaps

Cognitive film theorists have struggled with the issue of perception and interpretation, and in the next few paragraphs I shall provide a short summary of the most comprehensive position. In his seminal work *Making Meaning*, David

Bordwell comes up with terminology and different categories for comprehension and interpretation. Bordwell claims that “[c]omprehension is concerned with apparent, manifest, or direct meanings, while interpretation is concerned with revealing hidden, non-obvious meanings” (2). On the one hand, comprehension covers referential and explicit meanings, references the viewer extracts from the text that could be considered literal meanings (regarding the diegetic world, for instance, or such direct metaphors as the scales of justice). Interpretation, on the other hand, covers implicit and symptomatic meanings, meanings that the viewer derives from posing problems, asking questions, examining themes, or speculating about the intentionality of the author (8-9). In other words, comprehension refers generally to bottom-up processes of perception and identification and, to some degree, the high-order cognitive operation of causal reorganization of the *fabula*. But interpretation refers to high-order cognitive operations only, those postulating questions of intentionality, ideology, symbolic implications, irony, and other metatextual and contextual issues. In his book, Bordwell focuses on the latter rather than the former. “I will not be much concerned with comprehension,” he writes. “My stress here falls on interpretation, conceived as a cognitive activity taking place within particular institutions” (10). Much like some reception theorists (Staiger is a case in point), Bordwell here is nearly ready to discard the text in order to concentrate on a high-order cognitive activity that, according to him, operates in particular institutions that are independent of the text. But once he declares a possible theoretical separation between comprehension and interpretation and decides to focus on interpretation only, Bordwell (as I will show soon) loses ground.

Edward Branigan provides an adaptation of Alan Williams’s formulation that “when we watch a narrative film, we are actually watching four different films: a celluloid strip of material; a projected image with recorded sound; a coherent event in three dimensional space; and finally a story we remember (i.e., the film we think we have seen). There are perceptual ‘gaps’ between each of these four films in which certain facts are concealed and ‘forgotten’ about one film in order to perceive another” (84). Branigan here describes the phenomenon of ignoring editing within a scene—which implies spatial and often temporal skips—in favor of accepting the dramatic unity the scene conveys. Out of the same principles of dramatic coherence we ignore jump-cuts and other visual and audio inconsistencies that, according to this account, we perceive but dismiss for the overriding needs of the *fabula*. Branigan here accepts a modular approach to the viewer’s construction of narrative, one that fits nicely with his notion of narration since it evolves from hierarchical structural levels of both cognition, mind, and narrative.<sup>3</sup> Branigan claims that “comprehension proceeds by canceling and discarding data actually present, by revising and remaking what is given” (83). What is striking about his account is that it assumes that this process (comprehension) works in one direction, from top-down impositions (since the

needs of the narrative are computed in high-level cognitive procedures) onto the bottom-up perceptions. And while this account of comprehension is probably true much of the time, it should not be presented as the only possible one. In some cases, in, for instance, the cinema of Godard, jump cuts and other disruptions to cinematic conventions force (via bottom-up perception) a comprehension (or in Bordwellian terms an "implicit" interpretation) that the film is an essay about conventional cinema.

The opening shot of *Breathless* (Godard, 1959) shows a close-up on a page from a newspaper. A low and intimate voice-over of a man mutters some words, and the conventional assumption is that he is responding to what he reads on the page (i.e., that it is a point of view shot). But the camera then starts tilting up on the page, and continues up to show the face of Michel, revealing that we were seeing not the front but the back of the paper, thus canceling our hypothesis that it was a point of view shot. In other words, low-level perception of visual and audio information first encouraged a high-level hypothesis (regarding point-of-view shot) and then canceled it when Michel's face was revealed. To be sure, this moment is quick and probably does not produce a complete interpretation of theme, but it disorients the audience enough and, indicating that this film is not about to be a conventional film, offers a high-order implicit realization. In contrast, the last scene of a conventional mainstream film, *The Usual Suspects* (Brian Singer, 1995), changes the terms of narration and, thus, the relationship between film and viewer in its last few minutes. As many crime films do, throughout, the film encourages hypotheses about its events in conventional ways. But when it breaks and mocks these conventions in the final scene, it disables all the hypotheses produced so far, without offering another explanation as to what had actually happened. In that last scene, what is given, the information presented (in close-ups of objects, photos, text, etc.), is the source of the change in the state of all previous hypotheses, as everything that preceded (not only hypotheses, but events visually present as well) is marked as a fabrication.

Branigan's model of hierarchical comprehension leaves little room for such filmic interruptions in the perceiver's top-down operations, or at least these interruptions are always enveloped within high-order narratorial and perceptual structures. Moreover, if for Branigan the comprehension of existing data could be revised or readily written off, so much more is the danger with gaps or moments of indeterminacy that do not provide information at all. Although Branigan does not provide a comprehensive description of gap filling, he does refer to it occasionally:

By conceiving of narration as a type of verbal (and imagistic?) description offered by a spectator, one is, in effect, analyzing *interpretive* statements. One is mapping a course of thought, the use of language, rather than discovering the absolute properties of an object or discovering "cues" that are "in" an object—the text objectified. Interpretation thus construed exhibits something of the nature of an explanatory "theory." Interpretation in this sense includes the "filling in" of certain data (from the top-down) that seems to be "missing" at some moment in the text as well as the construction of macro-propositions

that are *about* the text though not strictly *in* it, or denoting it. Structures that are achieved in cognition cannot be reduced to a list of phenomenal forms or cues. We demonstrate our knowledge of narration, of "how to go on," by interpreting, by going on. (112)

Gaps, then, given that they are never "in" the text, are assumed to function only in high-order cognitive constructions of the text, ones that are filled "from the top-down." What I caution here against is the tendency—one both Bordwell and Branigan share—of separating top-down operations from bottom-up perceptions, and prioritizing the former without allowing for theoretical moments in which this dominance is broken.<sup>4</sup> As I have shown above with both *The Usual Suspects* and *Breathless*, the assumptions about top-down dominance leads to a partial and reductive description of what actually happens during viewing. But even more important than the examples that contradict this model of dominance, I would like to point out that some of the most intense moments in our cinematic experience are moments when a film surprises us from the bottom-up. In these moments, bottom-up perception interferes with "top-down" assumptions, forcing the viewer not only to reassess previous information but also to come up with new and revised logical models of causality and implications.

Gaps can be generally divided into two groups, implied and necessary. Implied gaps (also known as blanks) are gaps that are not essential to achieving dramatic coherence (as in assuming that the characters sleep, eat, and use the bathroom throughout the temporal duration of story time). Dramatically necessary gaps are gaps that necessitate inference production for the plot or text to cohere and make sense. Chaffin suggests that when we encounter a necessary gap we engage in producing bridging inferences, or what I earlier called gap filling. But the difference between implied inferences and necessary ones is striking. While implied inferences are processed as actually perceived information, necessary ones slow down comprehension while subjects search for plausible bridging mechanisms. In research on gap filling and comprehension, Havilland and Clark found that such sentences as "the murderer was one of John's friends" were read more slowly after "John died yesterday," but faster after "John was murdered yesterday." The conceptual gap between "death" and "murder" requires reorganization of previous data, and therefore slows down the comprehension of the text. More recent research shows that the slowed response has to do with memory activation. While we are engaged in reading, comprehension practices always attempt to maintain coherence on two levels: local, and global. Local information, on the one hand, is matched with immediately preceding information that is mostly available to short term memory on a scale of seven items, plus or minus two (see Miller). The maintenance of global coherence, on the other hand, requires the reader to compare the new information with previous information no longer available in working memory (see O'Brien et al.). While local coherence is important for mapping elements of the *syuzhet* as they are being perceived, global information is important for constructing the *fabula*.

In non-literary texts, at least two levels of representation of a text also operate at the same time, but these are called text-based representation and situation model. Whereas text-based representation is a representation of the text itself, a situation model is a representation of what the text is about: it contains both textual information and general knowledge of the world (O'Brien et al. 1200). In artistic texts, the situation model gives the global structure for the comprehension of the *fabula*. According to Paul van den Broek, cognitive psychologists have proposed that a situation model is established by processing the text into causal chains of inferences resulting in a perception that the text is coherent (423). To determine whether two events in the text exist in causal relations, van den Broek proposes four criteria:

According to the criterion of *temporal priority*, a cause never occurs after the consequence. According to the criterion of *operativity*, a cause is active when the consequence occurs. The *necessity in the circumstances* criterion reflects the fact that if the cause had not happened then the consequence would not have taken place, given the circumstances of the story. The *sufficiency in circumstances* criterion indicates that if the cause occurs, then the consequence will likely occur as well, given the circumstances of the story. (424-25)

Van den Broek cites much research that supports the existence of these criteria in postulating causal relations. Subjects seem to remember causally related events better than two unrelated events even when the two are adjacent in the surface structure (the *syuzhet*); in addition, dead-end events (ones that do not move the plot forward, and are not causally related to others) were forgotten faster than causally related events. Summing up much research from the 1980s, van den Broek claims: "Highly connected events are more often included in summaries [ . . . ], rated as more important [ . . . ], and retrieved more quickly [ . . . ] than events with few causal connections"(429). But the mental representation of the causal relations proposed or made available by the text is dependent on variable factors. Short-term memory capacity, for instance, may limit one's exploration of all possible connections while reading a focal event. Van den Broek proposes a Causal-Inference-Maker model (CIM) in which the causal criteria and the limitations determine the content and the types of inferences made, and hence form the conceptual and procedural *constraints* that operate on the inferential process (433).

Van den Broek's model, then, is based on the assumption that both the search for causal relations (according to the four conceptual criteria) and the procedural limitations on attention and memory guide the reading process and result in the situation model of the text. This model is a highly selective one that during perception prioritizes causal relations and claims that these relations affect the long-term status of the memory of the text. That is, events that do not seem as causally related (or potentially so) at the time of perception will be of a lower status in long term memory (the situation model of the text), and are less likely to alter the comprehension of the text when new information appears. The CIM postulates that

when adjacent events provide necessity and sufficiency, they will be connected as an inference. When the text provides a break, or a gap, the reader will search the memory for missing information and if it is found will reinstate an inference. If the text does not provide the necessary and sufficient events for an inference, the reader will fill in the gap by inferences based on world knowledge, a process van den Broek calls elaboration (434-35). The model is very useful for a theory of gap filling, but there are two problems with applying it to gap filling in films.

The first problem with van den Broek's model and its application to film has to do with his sole reliance on language:

It is important to note that several conditions need to be met in order for the reader to be able to construct a coherent mental representation of a text. [...] The function and meaning of words need to be identified and the words need to be combined to form a sentence or *proposition*. It is only after these tasks have been achieved that the reader can come to understand the relations among the individual events portrayed in the sentences. (442)

Van den Broek's model relies solely on language processing, and describes that process as following a linear and propositional formula, one that contains and defines the logical and causal relations among events. But while cinema incorporates natural languages, it communicates through auditory and visual images that provide multiple and often overwhelming amounts of detail at once. Such sensory information is processed simultaneously as well as serially, and is stored in memory both as propositional sets and as depictions (images). But when visual information is retrieved from memory, it is pulled up holistically. In other words, when the communication channel is largely visual, the basic processes van den Broek's model assumes are necessary for a causal memory-representation of texts do not always occur. The high-order restriction van den Broek places on the cognition of images is problematic. Images, as cognitive science shows us, are processed in a variety of ways, and stored in memory both as propositional sets and as descriptive images. Moreover, because films overwhelm us with a multiplicity of visual details, parts of the filmic image that may not seem relevant at the time of perception may become crucial for narrative comprehension later on. Indeed, if these visual elements are not stored in memory because they were not originally categorized as "meaningful," they will not be available for retrieval from memory at the time needed.<sup>5</sup>

The moth, in *Silence of the Lambs* (Jonathan Demme, 1991), for instance, seems irrelevant to the plot for much of the film's *syuzhet*, but when Clarice sees one on the serial killer's kitchen counter, it becomes a crucial narrative cue both for her and for the audience. That is, while perceivers assume that any detail (like the moth) in a thriller may turn out to be important, they cannot causally relate every detail at the moment perceived to the plot. Hence, according to van den Broek, they are less likely to remember it. But when we see the moth on the serial killer's kitchen counter, we easily and immediately remember and understand the narrative

importance of this cue. Unlike natural languages, which are cognitively processed by using high-order cognitive mechanisms (such as a lexicon), at least some parts of the visual image can (and often do) bypass the categorization and computation process. Although it is clear that the operations of visual perception and cognition are problematic for van den Broek's propositional sets, it is not impossible to articulate images into causal sets, but neither is it as easy and automatic as van den Broek suggests it is for reading. More important, propositional sets are not necessary, at least not as a pre-requisite, for the initial stage of perception and storage in memory of visual information. By prioritizing causal inferences, van den Broek's model leaves out much visual information that could not fit into propositional sets of necessity and sufficiency. But, clearly, because visual information is important for narrative comprehension, we need a more flexible account of narrative construction, one that allows nonpropositional sets to influence the situation model or the *fabula*.

The second problem with van den Broek's model lies in the restriction he places on long term memory, whereby only inferences that are made within the immediate cognitive constraints of memory capacity are stored and used when a gap occurs. In a series of experiments, O'Brien, et al. provided subjects with a series of sentences that posed a breakup in the coherence of the text. For instance, subjects read a text in which background information indicated that Mary is an avid vegetarian, but later on a sentence described Mary as ordering a cheeseburger. The information about Mary's being a vegetarian no longer exists in working memory, but it is pulled up and reactivated when subjects attempt to solve the problem of contradictory information.<sup>6</sup>

The results of these experiments show that more than causal information is considered for integration in the global and situation model. Instead, any and all memory-based variables can be reactivated and integrated not at the time of perception, but at the time they become relevant dramatically. Gaps then activate long term memory in an attempt to bring up all information that could be relevant for coming up with a coherent inference or a bridge.<sup>7</sup> In their research, O'Brien et al. show that information is stored in long-term memory even if at the time of perception it seems to be a dead end (information with no importance in a causal chain). In other words, this research challenges the hierarchical nature of memory storage and retrieval suggested by van den Broek. It is particularly important to our interpreting film, since film communicates via multiple channels of information, and these channels do not always tell the same story. The verbal causal relations may be different from the visual ones, and the cues of causality may be misleading altogether. For instance, *Silence of the Lambs* induces an inference that the FBI is at the house of the serial killer. Viewers draw this inference from the film's use of the conventions of parallel editing, its inclusion of a title indicating the FBI are in Calumet City, IL, and the auditory signal of the two door-bell rings that seem to connect the exterior of the killer's house with its interior. But these immediate



causal inferences are negated, first, when the serial killer opens the door to reveal not the FBI but Clarice (who is in Ohio) and is reiterated again, moments later, when the FBI breaks into an empty house in Illinois. A causal-inference model that restricts memory would make it very difficult to understand how Clarice could be at the door of the serial killer. But if other memories (taken as noncausal at the time) are available, the audience is able to correct the inference much faster and more efficiently. That Clarice just had an idea and was going to try to interview a dress maker, only so that she can collect "support evidence" for the trial, clearly becomes important when we see her at the door of the serial killer. The audience then realizes that what it thought was taking place in Illinois (at the basement of the serial killer) was actually happening in Ohio. The access to all background information in memory, as suggested by O'Brien et al., also resolves the problem of the reliance on language that was essential for van den Broek's model. If all long-term memory is accessible, then, whether it is retrieved propositionally or pictorially, it is available for gap filling and for the reorganization of the *fabula*, or the situation model.

In the case of a film, then, gaps activate both visual and verbal memories, and use both of them in a new integration of material aimed toward a reorganization of the *fabula*. I shall now turn to a detailed description of visual perception and cognition at work to understand how filmic gap filling operates in practice.

### A Gap in Action

Here, to illustrate gap filling, I will analyze in detail how viewers, in the production of inference and hypothesis, fill a gap in a scene from *Dead Poets Society* (Peter Weir, 1989). This analysis does not attempt to be conclusive or exhaustive, but I hope that it shows the merit of performing a cognitive analysis of gap filling. *Dead Poets Society* is a film that climaxes in a gap. The film focuses on the life of teenage boys in an upscale prep school, and particularly in the relationship they develop with a charismatic and inspiring teacher, Jack Eating (played by Robin Williams). One of the main characters is Neil Perry (played by Robert Sean Leonard), a good student from a modest-income home, whose real passion is theater. Neil excels in school, but is still forbidden by his rigid father to participate in a school play. Neil disobeys the father's orders, and is extremely successful as Puck (*A Midsummer Night's Dream*). But the father walks in during the play, and afterwards he takes Neil home and informs him he is enrolling the boy in an army academy to teach him discipline. In the following scene Neil commits suicide. But director Peter Weir never shows the suicide; instead, Weir shows us a series of shots leading to the suicide, and a series of shots after the suicide, but the film itself only confirms the suicide moments later. The film thus presents us with a gap, but it does support the hypothesis, or inference, regarding suicide that the audience develops. To explain how the gap is filled, I shall now examine the scene more closely.

The scene starts with a shot of Neil standing by the window, shirtless, lifting the window wide open. Eerie music fades in and plays throughout the scene until the moment of the suicide. Neil reaches out to grab the crown of thorns (part of his costume from the play), and puts it on his head. In the next shot, Neil lowers his hands and looks down. The camera lingers on Neil. While we know the wreath was a prop in the play, the shot creates an allusion to prototypical images of Christ as well. As metaphor, the crown of thorns does not necessarily operate in the surface structure of the text, since we can easily connect the shots by inferences based on the idea that the wreath is a symbol of the world of theater that Neil is about to lose. But once we realize (moments after the act) that Neil has killed himself, the allusion to Christ and the idea of martyrdom emerge quite naturally. Moreover, even if we don't consciously think of the Christ metaphor at the time of viewing, it still prepares us for the suicide in subliminal ways. That is, the image of Neil standing at the window with the wreath on his head is stored in memory not just as a section in a causal chain that articulates his lamentation over his theatrical career, but also as, potentially, a metaphor about martyrdom. The image does not necessarily foreshadow the suicide, but that it is readily available to support the suicide hypothesis is a condition that would be unlikely in van den Broek's model.

The scene continues with a series of close-up shots (3-9) on door-knob moving, door opening, feet on the floor, key, and a drawer opening, hands retrieving a wrapped object, and a pan shot from a dresser to a close-up shot of the father sleeping. All these shots are filmed in the dark of the night, and are obscured by the lack of light and by many of them being close-up shots. While the objects in the shots can be identified, it is unclear which space we are in at any given moment, whose feet and hands we see, or what is being retrieved from the locked desk drawer. That is, while the object-recognition aspect of image processing (also called the "ventral stream") is more or less readily available, placing these objects in relations to one another (in the "dorsal stream") is much more difficult to pin point, and the mapping of space, action, and who carries the action is at best obscure or ambiguous.<sup>8</sup> Moreover, the gun is never seen, as it is wrapped in cloth. The gun features in one shot of this sequence (shot 9), when it is being pulled out of the drawer, and the camera zooms out to reveal Neil sitting at a desk holding the wrapped object in the dark. But there is no clue at this point that the object is a gun, and there are no visual features that enable us to come up with a hypothesis that the cloth conceals a gun. This sequence of seven shots does not lend itself to a connectivity that can inform an inference. The shots seem somewhat disconnected (both spatially and dramatically), so rather than their moving the narrative forward, they, together with the eerie music, result in a sense of vague anticipation of some dramatic event rather than a clear observation that one is taking place.

In the next shot (#10) the father abruptly wakes up as if from a bad dream, and the music stops. He murmurs something about a sound, but given that we haven't heard anything, we first attribute his concern to a bad dream. Here we assume that

we have better means to make an inference than he does. We have not heard anything, and given that he was asleep, we assume that we are cognitively better informed (after all, we are awake . . .). But the father is haunted, and he proceeds to move through the house, turning on the lights everywhere he goes. His search (shots 11-22) is done in full light, and, unlike Neil's fragmented and dark journey, the father's is shot mostly in wide-angle shots giving us plenty of visual, contextualizing, information: Neil's room is empty, the wreath still on the open window sill, hallways are empty, and, finally, the office is clear and quiet, but we learn that the father smells something. While the previous sequence was full of ambiguous and difficult information, the father's sequence is visually available, but void of the narrative conclusion we anticipate. Finally, in shot 20, the father (now standing in the office) starts moving to his right, the camera cuts to his point of view shot, and, in a pan movement, it reveals a gun on the floor. The camera keeps panning right, as if the father is moving to the right, and an arm is seen on the floor. In the next shot, the father jumps forward, but the camera is shooting in slow motion, so his movement toward the desk is stretched in time. But we never see Neil on the floor, and we are only informed of his death in the following scene, in which the kids at the dorm are waking each other up with the news that Neil is dead. Still, the confirming word, "suicide," is mentioned only much later in the film, when Keaton (the charismatic teacher), is about to be fired.

The gap (the suicide itself) occurs between shot 9 (Neil sitting at the desk silhouetted) and shot 10, when the father wakes up abruptly. The filling in of the gap occurs sometime between shot 10 and shot 21 (in which we see the gun and hand on the floor). By the time shot 21 occurs, we are not surprised, but can (only partly) confirm the suspicion that Neil has shot himself. But given that this information (provided by our suicide hypothesis) is never provided in the surface structure of the text itself, our hypothesis originates from a complex, mostly high-order, operation performed on the information available.

Van den Broek suggests that when we process a text we try to connect adjacent events to one another. As I have shown above, the shots in the sequence leading to the gap do not connect well with one another, and the gap presents a real rupture, a break in the narrative flow. If necessity and sufficiency conditions (of cause and effect) cannot be met, van den Broek suggests that we search memory for missing information that will enable us to reinstate a causal relation and make a bridging inference. But the scene from *The Dead Poets Society* does not avail itself to those kinds of explicit causal relations either. Neil has not been suicidal throughout the film, there is no mention that the father owned a gun, and Neil's love of theater (and his father's anger at that) has not provided the main dramatic conflict of the film. There are thus no direct inferences that can be made based on a memory search. Moreover, the facts that Tom forbade Neil to participate in the play, that Neil disobeyed his father's orders, and that he is about to be enrolled in an army school could possibly lead to a hypothesis that Neil is leaving home, and that the wrapped

package is saved money. But, as I shall show shortly, there are reasons why this hypothesis is less likely to emerge.

Van den Broek suggests that in those cases where previous textual information is unavailable, we retreat to our naïve knowledge of the world, thus engaging in a process that he calls elaboration, a process that brings in extratextual information to create a bridge or to fill in the gap. Literary and film theorists discuss an intermediate stage between the long-term memory of a text and our general world-knowledge, one that relies on knowledge of the genre and of dramatic texts in general. Traditional drama, as identified by the Greeks, contains a set of dramatic conflicts that are usually resolved, but not before a climax occurs. In the sub-genre of prep-school films, a tragedy often occurs, and it is likely to revolve around the unfulfilled hopes and passions of one of the teenage protagonists of the film.

Using these generic norms while viewing the *Dead Poets Society*, audiences develop an expectation that something “bad” will happen to one of the protagonists. This vague expectation finds a home in the suicide hypothesis, and guides the process that leads to this choice as the preferred inference (over the “escaped home” hypothesis, for instance). But this account of gap filling still seems unsatisfactory. It at best provides us with a general direction for hypothesis production, but does not explain how over the span of 10 shots (at most) we have come to the conclusion that Neil killed himself. I believe that the specificity of the hypothesis has to do with a strong reliance on bottom-up perception of the filmic information in the scene.

The dramatic information in the scene is conveyed almost entirely by visual means (there are a couple of verbal exchanges between the father and mother, but they are not very informative, for the father doesn’t tell her what he thinks he heard or why he is up). This visual information is arranged (in terms of framing, shooting, and editing) to support the suicide hypothesis. The image of Neil as Puck/Christ sets the atmosphere of the scene. The disjointed series of dark and close shots that follows resists full cognitive processing, and therefore support the sense of looming danger. That we see the gun but cannot know that it is a gun (since it is concealed) prevents our connecting all the previous shots in a coherent causal chain, but instead gives the impression that the sequence is likely to culminate in a (yet unknown) climax. The visually ambiguous nature of the sequence prevents our anticipating what this climax is going to be, but both prepares us for its coming and ensures that, rather than surprised, we are ready to jump to conclusions. The abrupt end of music and the father’s jump punctuate the previous sequence, and as the search sequence begins, the audience realizes (from generic conventions) that the scene is bound to end in tragedy. It is then that the suicide hypothesis forms, and when we see the gun and the hand on the floor, it is nearly confirmed (“nearly” because Neil could still be just wounded).

The other major support for this hypothesis comes from the eerie music’s stopping at the moment of the suicide and returning when the father sees the gun.

In the first 9 shots, the music seems to indicate a looming danger or imminent tragedy. But when the music returns in shot #21, it becomes clear that it has stood in for Neil's mental state, his decision to kill himself, or for the suicide itself. I strongly believe that had the suicide hypothesis been solely a result of high-order cognitive activities, ones based mostly on causal sets of previously processed and categorized textual information and world knowledge, it would have been vaguer and would not have emerged so smoothly during the father's search sequence. But attentiveness to visual detail and the ability to refocus and reinterpret it when the narrative conditions change are both keys to the success of the gap-filling practices of this scene. What is required here is not just openness to bottom-up information, but an understanding that visual information is not processed cognitively in the same ways that language is. The concealed object, for instance, is categorized as such: an object that could be a gun, a wallet with money, a nostalgic theatrical idol, or any number of other objects. It is assumed that this visual memory will be important narratively, but it is not classified as part of a particular "propositional" (or, in Dretske's terminology, even "meaningful") object. It is stored as an ambiguous object, like many of the items in the dark shots preceding it, and is assumed that the information could become important and explicit at a later point. Visual reorganization works by our being able to retrieve images from memory, images that may not have been classified and categorized at the time of perception, but are being determined at the time of this later cognitive operation, the one of gap filling. The scene is effective (that is, not confusing), and elegant, precisely because high-order operations work in tandem with bottom-up perception.

The analysis of the gap in *The Dead Poets Society* shows that the activity of the perceiver is complex and depends on bottom-up perception of visual and auditory information, as well as on top-down assumptions about genre and the nature of drama in general. While interpretation requires such high-order cognitive operations as re-organization of story data in cause-and-effect structures, retrieval from memory of past events, speculation, and hypotheses production, it nevertheless operates in tandem with bottom-up perceptions. Moreover, the perceiver is always very sensitive and very ready to adjust interpretation based on the new flow of textual information. But cognitive and film theories of narrative tend to ignore or undermine the importance of bottom-up operations and, particularly, the specificity of perception and cognition of each of the different communicative tracks. As a result, existing models of filmic narrative construction are partial and need to be overhauled. In a longer book project, I hope to propose a more complete cognitive account of film narration and comprehension, one that will account for the perception and cognition of images, verbal information, sound effects and music and that will address the complex ways in which they interact with a perceiver to form the interpretive process of film.

## Notes

<sup>1</sup> The Russian Formalists drew the distinction between *fabula* and *syuzhet*. The *syuzhet* refers to the story events as they are organized in the text in a linear progression, though not necessarily in causal relations. The *fabula* is a construct, a reorganization of the *syuzhet* into a causal chain of story events in the right temporal order. The *fabula* then is a construct in the mind of the perceiver, a retelling of the narration into a coherent story.

<sup>2</sup> By high-order cognitive activities, I mean processes such as computation of data, memory retrieval, and problem solving. By low-order perceptual processes, I refer to attention to, and recognition of, actual information from the environment.

<sup>3</sup> Branigan here cites the work of various cognitive scientists such as Ray Jackendoff, Andy Clark, Jerry Fodor, Howard Gardner, and Marvin Minski. These researchers agree very little on the architectonics of the mind, but Branigan here does not adopt a particular model; rather, he is concerned with implying the importance of cognitive science to the understanding of the interpretive process of film, without getting into the different positions of these debates or without even making very concrete claims about how it operates in film comprehension.

<sup>4</sup> Bordwell claimed: "Interpreting (reading) is dissective, free of the text's temporality, and symbolic; it relies upon propositional language" (*Narration* 30).

<sup>5</sup> For a fuller account of the problem of visual perception, see Fred Dretske.

<sup>6</sup> Some experiments used qualifiers in the background information, as "Mary used to be a vegetarian" or "Mary is generally a vegetarian." These cases, too, slowed the reading response and showed that readers were activating long term memory in trying to solve a conflict with the term "vegetarian" that was already stored.

<sup>7</sup> In semiotic language, we call this a paradigmatic process (a vertical conjunction of information from different places in the text); the normal progression of the reading process is syntagmatic or linear.

<sup>8</sup> For a more detailed description of the ventral and dorsal stream, see Nakayama et al.

## Works Cited

- Bordwell, David. *Making Meaning: Inference and Rhetoric in the Interpretation of Cinema*. Cambridge, MA: Harvard UP, 1988.
- \_\_\_\_\_. *Narration in the Fiction Film*. Madison: U of Wisconsin P, 1985.
- Branigan, Edward. *Narrative Comprehension and Film*. New York: Routledge, 1992.
- Chaffin, R. "Knowledge of Language and Language About the World: A Reaction Time Study of Necessary and Invited inferences." *Cognitive Science* 3 (1979): 311-79.

- Dretske, Fred. "Meaningful Perception." Kosslyn and Osherson. 331-352
- Gee, James Paul, and François Grosjean. "Empirical Evidence for Narrative Structure." *Cognitive Science* 8 (1984): 59-85.
- Haviland, S. E., and H. H. Clark "What's New? Acquiring New Information as a Process in Comprehension." *Journal of Verbal Learning and Verbal Behavior* 13 (1974): 512-21.
- Kosslyn, Stephen M. and Daniel N. Osherson, eds. *An Invitation to Cognitive Science: Visual Cognition*. Vol. 2. Cambridge, MA: Massachusetts Institute of Technology P, 1995.
- Miller, George A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *The Psychological Review* 63 (1956): 81-97.
- Nakayama, Ken, J. He Zijiang, and Shinsuke Shimojo. "Visual Surface Representation: A Critical Link Between Lower Level and Higher Level Vision." 1-70.
- O'Brien, E. J., M. L. Rizzella, J. E. Albrecht, and J. G. Halleran. "Updating a Situation Model: A Memory Based View." *Journal of Experimental Psychology: Learning, Memory and Cognition* 24.5 (1998): 1200-10.
- Staiger, Janet. *Interpreting Films: Studies in the Historical Reception of American Cinema*. Princeton, NJ: Princeton UP, 1992.
- van den Broek, Paul. "The Causal Inference Maker: Towards a Process Model of Inference Generation in Text Comprehension." *Comprehension Processes in Reading*. Ed. D. A. Balota, G. B. Flores d'Arcais, and K. Rayner. Hillsdale, NY: Laurence Erlbaum Associates, 1990. 425-45.

